

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6
Run on:	January 8, 2004, 10:25:26 ; Search time 48 Seconds
Scoring table:	BLOSUM62
Gapop	10.0 , Gapext 0.5
Searched:	1107863 seqs, 15876573 residues
Minimum DB seq length:	0
Maximum DB seq length:	200000000
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries
Database :	A_Geneseq_19Jun03:*
1:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1980 DAT:*
2:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1981 DAT:*
3:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1982 DAT:*
4:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1983 DAT:*
5:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1984 DAT:*
6:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1985 DAT:*
7:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1986 DAT:*
8:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1987 DAT:*
9:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1988 DAT:*
10:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1989 DAT:*
11:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1990 DAT:*
12:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1991 DAT:*
13:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1992 DAT:*
14:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1993 DAT:*
15:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1994 DAT:*
16:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1995 DAT:*
17:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1996 DAT:*
18:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1997 DAT:*
19:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1998 DAT:*
20:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA1999 DAT:*
21:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA2000 DAT:*
22:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA2001 DAT:*
23:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA2002 DAT:*
24:	/SIDS1/gcdata/geneseq/geneseqp-emb1/AA2003 DAT:*
SUMMARIES	
Result No.	Score
1	3557
2	3498
3	3030
4	2285
5	1982
6	1982
7	1917
8	1842.5
9	1228.5
10	34.5
DESCRIPTION	
Query	Match Length DB ID
	Amino acid sequence
	Human polypeptide
	Human polynucleotide
	Human protein seq
	Novel human diagno
	Novel human diagno
	Novel human diagno
	Drosophila melanog
	Novel human diagno
ALIGNMENTS	
RESULT 1	
ID	AAB08732 standard; Protein: 705 AA.
AC	AAB08732;
XX	DT 02-JAN-2001 (first entry)
XX	Amino acid sequence of a human OLD-35 polypeptide.
XX	OLD-35; OLD-64; OLD-137; OLD-139; OLD-142; OLD-175; cancerous phenotype; cellular senescence; terminal differentiation; growth suppression; aging process; type I interferon; cancer cell; tissue regeneration; ss.
XX	OS Homo sapiens.
XX	WO200046591-A2.
XX	PD 10-AUG-2000.
XX	PP 02-FEB-2000; 2000WO-US02920.
XX	PR 02-FEB-1999; 99US-0243277.
XX	PA (UYCO) UNIV COLUMBIA NEW YORK.
XX	Fisher PB, Leszczyniecka M;
XX	WPI: 2000-532905/4B.
XX	DR N-PSDB; AAA64608.
XX	PT Novel isolated nucleic acid encoding an OLD-35 or OLD-64 protein useful

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Db	1	DGPFLIPRRDRALTLQVRAILWSSA-SARRAVADLGNKL
LSSGKLARPADSAVWGGD	77
QY	61	TAVMTAVSKTKPSSQFMPFLVUDYPROKAMAGRIPINMURREVGSDEKLTTSRIDS	120
Db	78	TAVMTAVSKTKPSSQFMPFLVUDYPROKAMAGRIPINMURREVGSDEKLTTSRIDS	137
QY	121	IRPLPEAGYVYDITQTLCNLIAVDGVMNPDVLAINGASVALSISDIPNGPVGAVRIGID	180
Db	138	IRPLPEAGYVYDITQTLCNLIAVDGVMNPDVLAINGASVALSISDIPNGPVGAVRIGID	197
QY	181	GRYVNPTRGMSSTNLWVAGAPKSIQVMLEASARNILQODFCIAKVGKTYOPIO	240
Db	198	GRYVNPTRGMSSTNLWVAGAPKSIQVMLEASARNILQODFCIAKVGKTYOPIO	257
QY	241	GIQIQVKEKGTYKTRTPQKLFIPSPEVKYIKLAMERLYAVFTDYSHKVSRDAVKIR	300
Db	258	CIQDVKETGTYKTRTPQKLFIPSPEVKYIKLAMERLYAVFTDYSHKVSRDAVKIR	317
QY	301	LDTBESQKKEPKPREADPYETTESPNVAKBVERSTIVANBYKRCGDRGDTSLRANSCEVDMF	360
Db	318	LDTBESQKKEPKPREADPYETTESPNVAKBVERSTIVANBYKRCGDRGDTSLRANSCEVDMF	377
QY	361	KTLGGSALFORGOTOTLCTYFDSLESGIKSDQVITAINGIKDKNPLHYFPPATNEI	420
Db	378	KTLGGSALFORGOTOTLCTYFDSLESGIKSDQVITAINGIKDKNPLHYFPPATNEI	437
QY	421	GRVTLGILNRRELGHGALAKALYKPVIPDFPTITRVTSEVLESNGSSMASACGGSLALMD	480
Db	438	GRVTLGILNRRELGHGALAKALYKPVIPDFPTITRVTSEVLESNGSSMASACGGSLALMD	497
QY	481	SGVRISSAVAGVATGIVLTKDPEKSERIDYRLTIDILGIDYNGMDPKLAGTNKGITAL	540
Db	498	SGVRISSAVAGVATGIVLTKDPEKSERIDYRLTIDILGIDYNGMDPKLAGTNKGITAL	557
QY	541	QADIKLPGIPKIVHAIQDSVAKKEILQIMNKTKPSRASKRKNPWPVETVQWLSKR	600
Db	558	QADIKLPGIPKIVHAIQDSVAKKEILQIMNKTKPSRASKRKNPWPVETVQWLSKR	617
QY	601	AKFVPGGGNLUKKQAEGTGVTISQDDETSVFAFTPSVMEARDPITECKDDPQOLE	660
Db	618	AKFVPGGGNLUKKQAEGTGVTISQDDETSVFAFTPSVMEARDPITECKDDPQOLE	677
QY	661	FGAVYTATTEIRDGVWVLYPNMTAVLHNTQDNLRL	700
Db	678	FGAVYTATTEIRDGVWVLYPNMTAVLHNTQDNLRL	717
RESULT 3			
ID	ABG30875	standard; Protein; 675 AA.	
XX			
AC	ABG30875;		
XX			
DT	21-OCT-2002	(first entry)	
DB	Human polynucleotide phosphorylase	74.25.	
KW	Human; enzyme; polynucleotide phosphorylase	74.25; malignant tumour;	
KW	haemopathy; human immunodeficiency virus infection; HIV;		
KW	immunological disease; inflammation.		
OS	Homosapiens.		
XX			
PN	CN1341720-A.		
XX			
PR	27-MAR-2002.		
XX			
PP	05-SEP-2000; 2000CN-0119892.		
XX			
PR	05-SEP-2000; 2000CN-0119892.		
XX			

Db 573 VYATITIEIRDTGVMKLYPMNTAVLHNQDQRKI 609

Db 18 DGPFLPLPRRDRALTQOLTRALMSAGSRAVAVDGLNRKLBISSGKLARFADGSAAVUSGD 77

RESULT 4

Db 61 TAVMTATSKTKPSQSPMLVUDYRQRAAGRIPMTYLRQEVGTSKELTSRIDS 120

AAB92684

ID AAB92684 standard; Protein: 504 AA.

Db 78 TAVVTTAVSKTKPSQSPMLVUDYRQRAAGRIPMTYLRQEVGTSKELTSRIDS 137

XX

AC AAB92684;

XX

DT 26-JUN-2001 (first entry)

XX

DB Human protein sequence SEQ ID NO:11065.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EPI074617-A2.

XX

PD 07-FEB-2001.

XX

PP 28-JUL-2000; 2000BP-0116126.

XX

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-030253.

PR 11-JAN-2000; 2000JP-018776.

PR 02-MAY-2000; 2000JP-018767.

PR 09-JUN-2000; 2000JP-0241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR XX

XX

WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

XX

PS Claim 8; SEQ ID 11065; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the combination of oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAB03166 to AAB13628 and AAB13633 to AAB1842 represent human cDNA sequences; AAB2446 to AAB95893 represent human amino acid sequences; and AAB13629 to AAB13632 represent oligonucleotides, all of which are used in the exemplification XX

Sequence 504 AA;

Query Match 64.2%; Score 2285; DB 22; Length 504;

Best Local Similarity 100.0%; Pred. No. 1_9e-192;

Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ

1 DGPFLPLPRRDRALTQOLTRALMSAGSRAVAVDGLNRKLBISSGKLARFADGSAAVUSGD 60

Db 301 LDTEHQKKEPKPEADPITEESNVAKEVFRASIVLNEXYKRCGDRDLSLRNSCEVNDP 360

Db 318 LDTEEQKKEPKPEADPITEESNVAKEVFRASIVLNEXYKRCGDRDLSLRNSCEVNDP 377

RESULT 5

Db 61 ABG08546 standard; Protein: 899 AA.

XX

AC ABG08546;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #8537.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200115067-A2.

XX

PD 11-OCT-2001.

XX

PP 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

XX

DR AAS2733.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess PT biodiversity -

XX

PS Claim 20; SEQ ID No 38905; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences, (I) is useful as hybridisation probes, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377, represent novel human diagnostic amino acid sequences. ABG0010-ABG30377, represent novel human specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

SQ

Sequence 899 AA;

Query Match 55.7%; Score 1982; DB 22; Length 899; Best Local Similarity 73.2%; Pred. No. 2.8e-165; Matches 408; Conservative 18; Mismatches 29; Indels 102; Gaps 2;

Qy 10 DRAFTQVLQRALMSSAGSRAVAVDGLGNRKLTSGGKLLRADFGSAVQSGDTAVNTAVS 69
Db 440 DRAILQQLQVRAALISSTSGSQAVALDGLGNRKLTSGGKLLRADFGSAVQSGDTAVNTAVS 499

Qy 70 KTKPSISQFMPFLVVDYRQKAAAGRIPNTYVLRREVGTSDKEILTSRIDSIRPLFPAGY 129
Db 500 KTKPSISQFMPFLVVDYRQKAAAGRIPNTYVLRREVGTSDKEILTSRIDSIRPLFPAGY 521

Qy 130 FYDTQVLCLNLAVDGTNEPDVLAINGASVALSLSDIPWNGPGVAVRIGIDGEGYVNPTR 189
Db 522 ----- VGMTDGECVNPTR 535

Qy 190 KEMSSSTINLVAGAPSKQVIMLEASAEINTQODFCHAIKGVKTTQOIQIGIQLVKET 249
Db 536 KEMSSSTINLVAGAPSKQVIMLEASAEINTQODFCHAIKGVKTTQOIQIGIQLVKET 595

Qy 250 GVTKTRPQLTIPSPEVTKYTHKLAMERLYVAFPTDYEHDKYSRDEAVNKIRLDTRBQKE 309
Db 596 GVTKTRPQLTIPSPEVTKYTHKLAMERLYVAFPTDYEHDKYSRDEAVNKIRLDTRBQKE 655

Qy 310 KPEADAPYEIESFNTVAKFVSTVNEVYGRCPDPLTSRANVSEVDMKTRLGSALP 369
Db 656 KPEADAPYEIESFNTVAKFVSTVNEVYGRCPDPLTSRANVSEVDMKTRLGSALP 715

Qy 370 ORGQTOVLCTTFSLESGINSQDVITAISIKDQKFMHLHEPPVATNEIGKVTLNRR 429
Db 716 ORGQTOVLCTTFSLESGINSQDVITAISIKDQKFMHLHEPPVATNEIGKVTLNRR 775

Qy 430 ELGHGALAKALYKPVTPRDFPTIRTVSEVERSSSSMASAGGSLALMDSGVRISSAV 489
Db 776 ELGPGLAELAKALYKPVTPRDFPTIRTVSEVERSSSSMASAGGSLALMDSGVRISSAV 818

Qy 490 AGVAGLGLVTKDPEKEKIRDPLTULGIRDYNGMDPKAGTNGITLAQADIKLPGI 549
Db 819 :GVAMGLATKTDKEKIRDPLTULGIRDYNGMDPKAGTNGITLAQADIKLPGI 877

Qy 550 PIKIVMEAQOASVAKK 566
Db 878 TMKIVMEAQOASVAKR 894

RESULT 6
ABG17275
ID ABG17275 Standard; Protein; 899 AA.
XX
AC ABG17275;
XX
DT 18-FBB-2002 (first entry)

XX Novel human diagnostic protein #17266.
DE
KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
KW Homo sapiens.
XX
DN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PR 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSEQ) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT,
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS81462.

PT
PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
XX
PS Claim 20; SEQ ID NO 47634; 103PP; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377, represent novel human diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

SQ

Sequence 899 AA;

Query Match 55.7%; Score 1982; DB 22; Length 899; Best Local Similarity 73.2%; Pred. No. 2.8e-165; Matches 408; Conservative 18; Mismatches 29; Indels 102; Gaps 2;

Qy 10 DRAFTQVLQRALMSSAGSRAVAVDGLGNRKLTSGGKLLRADFGSAVQSGDTAVNTAVS 69
Db 440 DRAILQQLQVRAALISSTSGSQAVALDGLGNRKLTSGGKLLRADFGSAVQSGDTAVNTAVS 499

Qy 70 KTKPSISQFMPFLVVDYRQKAAAGRIPNTYVLRREVGTSDKEILTSRIDSIRPLFPAGY 129
Db 500 KTKPSISQFMPFLVVDYRQKAAAGRIPNTYVLRREVGTSDKEILTSRIDSIRPLFPAGY 521

Qy 130 FYDTQVLCLNLAVDGTNEPDVLAINGASVALSLSDIPWNGPGVAVRIGIDGEGYVNPTR 189
Db 522 ----- VGMTDGECVNPTR 535

Qy 190 KEMSSSTINLVAGAPSKQVIMLEASAEINTQODFCHAIKGVKTTQOIQIGIQLVKET 249

Db 536 KEMSSSTLNWVAGPKSQVMLSEAENILQDQFCHAIVGVKTQQ1QGQLQVKI 595
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represents novel human
 CC diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Db 596 GVTKTPQKLFPTPSPEIVKTYTHKLMERLYAVFTDYEHDKVSRDEAVNKIRLDRTEOLKE 309
 CC
 Db 596 GVTKTPQKLFPTPSPEIVKTYTHKLMERLYAVFTDYEHDKVSRDEAVNKIRLDRTEOLKE 655
 CC
 Db 310 KFPEADPYETIESNVAKEVFRSTVNLNEYKRCGRDLSLRNVSCEUDMFKTHGSAF 369
 CC
 Db 656 IPEVDVLYEIESFNTVAKEVFRSTVNLNEYKRCGRDLSLRNVSCEUDMFKTHGSAF 715
 CC
 Qy 370 QRGQTVLCTVTFDSLESGIKSQDQVITATINGIKOKNFMHYEFPYATNEI 429
 CC
 Db 716 QRGQTVLCTVTFDSLESGIKSQDQVITATINGIKOKNFMHYEFPYATNEI 775
 CC
 Qy 430 ELGPGALAEKALYKPVIPKDFKAGTNGKITALQADIKLPGI 489
 CC
 Db 490 AGVAGLGLVTKDPEGEIEYDYLTDIGEYDNGMDFKAGTNGKITALQADIKLPGI 549
 CC
 Db 819 -GVAMGLATKNDLKEBIEDYHLLTDILGIEAVNGMDFKAGTNGKITALQADIKLPGI 877
 CC
 Qy 550 PIKIVMBAIQQASVAKR 566
 CC
 Db 878 TMKIVMEATQQASVAKR 894
 CC
 RESULT 7
 ABG17276 ID ABG17276 standard; Protein; 439 AA.
 XX
 AC ABG17276;
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #17267.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0549217.
 PR 25-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR MPI; 2001-639362/73.
 DR N-PSDB; ARS81463.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 47635; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or

Db 676 ELGPGALAEKALYKPVIPKDFKAGTNGKITALQADIKLPGI 818
 CC
 Qy 490 AGVAGLGLVTKDPEGEIEYDYLTDIGEYDNGMDFKAGTNGKITALQADIKLPGI 877
 CC
 Db 819 -GVAMGLATKNDLKEBIEDYHLLTDILGIEAVNGMDFKAGTNGKITALQADIKLPGI 877
 CC
 Qy 550 PIKIVMBAIQQASVAKR 566
 CC
 Db 878 TMKIVMEATQQASVAKR 894
 CC
 RESULT 7
 ABG17276 ID ABG17276 standard; Protein; 439 AA.
 XX
 AC ABG17276;
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #17267.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0549217.
 PR 25-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR MPI; 2001-639362/73.
 DR N-PSDB; ARS81463.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 47635; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or

RESULT 8
 ABG58546 ID ABG58546 standard; Protein; 748 AA.
 XX
 AC ABG58546;
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 2430.
 XX Drosophila melanogaster developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS
 PN WO200171042-A2.

27-SEP-2001.
PD
XX
PR
XX
PR 23-MAR-2001; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-65686/75.
XX DR N-FSDB; ABL02649.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -
XX Disclosure; SEQ ID NO 2430; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU6176-ABU30511), expressed DNA sequences (ABU1840-ABU16175) and the encoded proteins (ABB7737-ABB72072).
CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX Sequence 748 AA;
SQ

Db	508	AIGUTKENDDTGHQLQYRILTDLGI	EDYMGMD	-----M	544
Qy	552	KIVMEAIQASVAKKEKIQIOMNIKTSKPRASKRKGNGPQVUTVQVPLSKRAKFVGGYNYL	611		
Db	545	KVVMBSLOKATDANSILDIMSEARPRPKPQKBSMPVSBTLLVTPQORALIGPSGLHM	604		
Qy	612	KKUQAEQTYTISQVDEEFTPSVPAFTPSVHHEARDEITEICKUDQQQLEQGAVTATITE	671		
Db	605	KRIVYETGTSLTAVIDEHNVPAPSQAMDEAKSLIEGTMVKERVDPLEGGIYVAKTE	664		
Qy	672	IRDTGVMVLYKPNMTAVLHNTQDNLRL	700		
Db	665	LRDTGVMVLYKPSMPALLHNSQDORKI	693		
RESULT 9					
ABG08547					
ID	ABG08547	standard; Protein; 541 AA.			
XX					
AC	ABG08547;				
XX					
DT	13-FEB-2002	(first entry)			
XX					
DE	Novel human diagnostic protein #8538.				
XX					
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200175067-A2.				
XX					
PD	11-OCT-2001.				
XX					

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 541 AA;

Query Match 34.5%; Score 1228.5; DB 22; Length 541;

Best Local Similarity 51.7%; Pred. No. 3.7e-99; Matches 306; Conservative 48; Mismatches 133; Indels 105; Gaps 18; DB 6 TNEIGKVUTGLNRLRELGHGALAKALYKPVIPRDPFPFTIRVSEVLESNGR----- 55

QY 97 TNYLRLREVGTSKDEILTSRTRPLPFPAGYFDTQVLCNLALAVDGYNEDVLAINGA 156
 QY 56 -----RIG-----OPPASRPRVYRAPLPPPLVRAPI 86

QY 211 MLEASABENILOQDFCHAIKVQYQI--ILOGIQLVKETGVTKRTPQKLFITSPEIVK 210
 QY 87 ASRSPGSEBVCTVTP--DSLSSGGIKSPQVITNGIK-----DKNP----- 124

QY 269 YTHKLAMERIYAVFTYDYEHDKVRSDAENKIRLD---TEELQKFKPFPADPYETESFN 324
 QY 125 -----MLHVEFPPIATNEIGKVUTGLNRLRELGHGALAKALYKPVIPRDPFPFT----- 171

QY 325 WAKEVERFSTVNLNEYKRCDFDRDLS--LRWNSC---VDMFKTULHGALFQGQTOU 377
 QY 172 RVTSEVLES--NGMERRIGQGPASSRSPVVRAPLPLVRAPI 224

QY 378 CTVTDLSLEGIKSQVITINGIKDKNFMHLHYEPYPAATNEIGKVUTGLNRLRELGHGALA 437

QY 225 CTVTDLSLEGIKSQVITINGIKDKNFMHLHYEPYPAATNEIGKVUTGLNRLRELGHGALA 284

QY 438 EKALYKPVIPRDPFPFTIRVSEVLESNGSSMASACGGSALMDSCVPISSAVAG-VAIGL 496
 QY 285 EKALYKPVIPRDPFPFTIRVSEVLESNGSSMASACGGSALMDSCVPISSAVAGRTNWD 344

QY 497 VTKTDPEBKGIEIYDYLTDI-LGIEDYNGMDPKLAGTNKGITALQADIKLPGIPKTM 555
 QY 345 SPKDPDEBKGIEIYDYLTDI-LGIEDYNGMDPKLAGTNKGITALQADIKLPGIPKTM 404

DB 556 RAIQQASVAKKEILOQMNKTSKPRASKRKGWPVETVQVPLSKRAFKVGGCNLKKQ 615
 DB 405 EAQDQASVAKKEILOQMNKTSKPRASKRKGWPVETVQVPLSKRAFKVGGCNLKKQ 464

QY 616 AETGYTISQVDETF-PSVFAFPATPSYHARDF-ITEICKDQEOOL-EFGAV 664
 QY 465 RETGVTISQVDEETVPCNLHQHVPVMLEGKRLSFEICKDQEOAARNLGAV 516

RESULT 10
 ID AAU33947 standard; Protein; 696 AA.

XX AC AAU33947;
 XX DT 14-FEB-2002 (first entry)

XX DB Staphylococcus aureus cellular proliferation protein #223.

XX KW Antisense; prokaryotic cellular proliferation protein;
 XX KW antibiotic; antibacterial; drug design.

XX OS Staphylococcus aureus.

PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-200848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-233625P.
 PR 22-DEC-2000; 2000US-227931P.
 PR 16-FEB-2001; 2001US-249308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr CJ;
 PI Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 DR N-FSDB; AASS1806.

XX New Polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX Example 3; Seq ID No 5443; 511P; English.

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 696 AA;

Query Match 33.5%; Score 1191.5; DB 22; Length 696;

Best Local Similarity 39.2%; Pred. No. 1e-95; Matches 274; Conservative 117; Mismatches 249; Indels 59; Gaps 12; DB 2 SQRKVFKTEWAGRSLTETGQLAQANGAVLRYGDTVVLSTASKEPRDGFPTV 61

QY 84 DYRQKAAGAAGRIPNLYRREVGSDKEILTSRTRPLPFPAGYFDTQVLCNLALAVD 143
 DB 62 NYEERMYAKGKIPGPGKFKREGPGCDATITLDRPDRPLPFPKGYKADQVQINMVISAD 121

QY 144 GVNEDPVLAILNGAVALSLSDIWNGPGAVRIGIODEYVNPTRKEMSSSTLNUVAG 203
 DB 122 PDCSPOMAMIGSMSALSVSDIPFQGPAGVNVYIDCKYINPTVSKESVSLDLEVAG 181

QY 204 APKSQIVMLEASABENILOQDFCHAIKV3-----VKTQOIQIGIQLVKETGVTKRTP 256
 DB 182 -HKDAVNVWEAGASBETIQMLEAIFPSHBEIQRLVLPQQQVHDQVQKQ----- 232

QY 257 OKLUTPS-----PEIYKTYTHKLAMELYAVFTYDYEHDKVRSDAENKIRLDTEEOKKE 309
 DB 233 ---FIPABERDEALVERIKSLTEERGLKETVLTF-----DKQDRENLNLK---EEIVNE 281

QY 310 KFPEADP-----YEIISFNVVAKERVSIVLNERYKCDGRDLTSJRNVSCEVDMFKTL 363
 DB 282 FIBEDPNELLKEVVAIINELVKEERLIADEKIPRDPGRKPDTRPLDSBVGILRT 341

QY 364 HGSALFQGQTOQVQCTVTFDLSLSRGKIKSDQVTAINGIKDKNFMHLHYEPYPAATNEIGKV 423
 DB 342 HGSGLFTRQGQTOAISVLTIGAL---GRYQLIDCLGPBEEKRFMHHTNPPNFSVGECPOV 397

QY 424 TGLNRRELGHGALAEKALYKPV--RDPFPFTIRVSEVLESNGSSMASACGGSALMDS 481

```

Db      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 457
398 RAPGRREIGHGALGERALKVIIPTDADPRTIRIVSEVLENSGSSQASICSTLALMDA
QY      ||||: GVPISAVAGAIGLVKTPDPERGELEYRFLTDLGLGFDKNGDMKIAKGNGKITALQ 541
482 GVPISAVAGAIGLVKTPDPERGELEYRFLTDLGLGFDKNGDMKIAKGNGKITALQ
Db      ||||: GVPIKAPVAGIAMLVRED-----SYTILTDIQMGEDALGDMDFKVGAKREGITAO 510
458 ADIKURGIPIKIVMARIQASVAKKEILQIMKNTISKPRASKRKGKPVETVQPLSKRA
QY      542 ADIKURGIPIKIVMARIQASVAKKEILQIMKNTISKPRASKRKGKPVETVQPLSKRA 601
511 MDIKIDGUTBTERLEALQDGRGRBLIMNMQTIDPRTLSAYAKVUWIKPDKIR
Db      602 KFVPGGGYNLKKLQAETQVPMKLYPNTAVLLNTQNLRL 700
QY      662 GAVYATITERDGTGWMKLYPNTAVLLNTQNLRL 700
Db      626 GQTYQATVKRIEKYGAFVGLFPGKDA-LIHSQSKRRI 663

```

```

RESULT 11
AAU36732 standard; Protein: 698 AA.
ID AAU36732;
XX AC AAU36732;
XX DT 14-FEB-2002 (first entry)
XX DE Staphylococcus aureus cellular proliferation protein #902.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-20727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELTR-) ELTRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
XX DR WPI: 2001-611495/70.
XX DR N-PSDB; A0554591.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
```

CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 Note: The sequence data for this patient did not form part
 CC of the printed specification, but was obtained in electronic
 format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.
 Sequence 698 AA;

Query Match 33.5%; Score 1191.5; DB 22; Length 698;
 Best Local Similarity 39.2%; Pred. No. 1e-95; Indels 59; Gaps 12;
 Matches 274; Conservative 117; Mismatches 249; Indels 59; Gaps 12;

```

Db      24 SAGSRAVADLGNRKLEISSKKLARRDGSAGWVQSDGATVNTAVSKTKFSPSQFMLV 83
QY      84 DYRKQAAAGRIPNTYRREVGTSDKEDILSRISIDSIRPLPAGYFYTDTQLCNLAUD 143
Db      62 NYEERKTYAAGKIPGGFKKREGRPDDAFLTARLDRPFLPKGKHDVQMMVLSAD 121
Db      144 GYNERPDYLAINGASVAVLSSLSPWMCPPGAVRIGIDGEVYVNPTKEMSSTLNIVAG 203
Db      122 PDCSPONAMIGSSMALSVPDIFCGPDIAGVNVGVDGKYINPNTVEKEVSRDLEVAG 181
Db      204 APKSQITVLEASAENILQDQDCHAIKG-----VKTQOIQ3IQOLVKETGVTKRTP 256
Db      182 -HKDADNNVNEGASETTEQMLEAFFFHHBIRLAFQDQDPLHOPQDUDHOPVKE----- 232
Db      257 OKLFTPS-----PBTVKYTHKLAMERLYAVFTDYEHKVSDEANKURLDTBQOLKE 309
Db      233 ---FIPADEALVERPKISLTLEEKLKEUTP---DKQDRENIDNUK--BIVNE 281
QY      310 KFPEADP-----YELTSFNVAKVERFSIYLNEXYKRCGDRDLTSLRNTSCEVDMFKTL 363
Db      282 FIDEEDEPENELIKEVYVAILNELVKEVRRLIADERKIPDKRKPDBIRPLDSEVGLPLRT 341
Db      364 HGSALQRGQVQVLTIVTFFSLESQSIKSQDQVTAINGIKONKMPHYEFFPYATWIGKV 423
Db      342 HGSGLTRGOTQALSYTILGAL---GDXOLGSLGPEBEKRPHMYNPFNSVGETGPV 397
Db      424 TGLNRDEBLGIGALAKSVALYVIP--RDFPTIRVSEVLENSGSSMASACGGSLALMDS 481
Db      398 RAPGRREIGHGALGERALKVIIPTDADPRTIRIVSEVLENSGSSQASICSTLALMDA 457
QY      482 GVPISAVAGAIGLVKTPDPERGELEYRFLTDLGLGFDKNGDMKIAKGNGKITALQ 541
Db      458 GVPIKAPVAGIAMLVRED-----SYTILTDIQMGEDALGDMDFKVGAKREGITAO 510
Db      511 MDIKIDGUTBTERLEALQDGRGRBLIMNMQTIDPRTLSAYAKVUWIKPDKIR
QY      602 KFVPGGGYNLKKLQAETQVPMKLYPNTAVLLNTQNLRL 700
Db      662 GAVYATITERDGTGWMKLYPNTAVLLNTQNLRL 700
Db      626 GQTYQATVKRIEKYGAFVGLFPGKDA-LIHSQSKRRI 663

```

RESULT 12
ABP40063 ID ABP40063 standard; Protein: 706 AA.
XX AC ABP40063;
XX DT 24-JUL-2002 (first entry)
XX DB Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4908.

XX	KW	Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX	KW	antibacterial; gene therapy.
OS	Staphylococcus epidermidis.	
XX	US380370-B1.	
XX	PD	30-APR-2002.
XX	PD	13-AUG-1998; 98US-0134001.
XX	PR	14-AUG-1997; 97US-05579P.
XX	PR	08-NOV-1997; 97US-064964P.
PA	(GENO-) GENOME THERAPEUTICS CORP.	
XX	PI	Doucette-Stamm LA, Bush D;
XX	DR	WPI: 2002-381555/41.
DR	N-PSDB; ABN9208.	
XX	PT	Novel isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections - Disclosure; SEQ ID 4908; 267pp; English.
XX	CC	AB90538 to ABN93374 represent Staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP5524 to ABP5690. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection.
CC	CC	N. B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site.
CC	SQ	Sequence 706 AA;
Query Match	32.8%	Score 1165.5; DB 23; Length 706;
Best Local Similarity	38.7%	Pred. No. 2.1e-03; ID
Matches	270;	Conservative
CC	CC	Mismatches 123; Indels 55; Gaps 13;
QY	24	SAGSRAVAVDGLGNRKRUEISSGKLRADAGSAVQSGDTAVMVTAVSKTKPSPSQFPLVV 83
Db	7	SQEKKVKTEWAGRSLTIEQQLAKOANGAVLVLVKGDTVVLSTATASKEPRGDFFLTV 66
QY	84	DYRQKAAAGKRIPTYLRLREVGTSDKELTSRISDRIPRPGFPRYDPTQYDTCNLIAVD 143
Db	67	NYEKVVAAGKIPGPKREGRPDEATLRLDPRIPRPLPKYRHDVQMNIVLSD 126
QY	144	GVNEEDVLAINGASVSLSDIPWNGPGAVRIGIDGEYVNPTRKEMSSSTLNLVAG 203
Db	127	PDCSEMAAMIGSSMALSVDIPFQDPLAGYVNGVYDGKVVINPSVADKRSRDLVAG 186
QY	204	APKSDQVMLEAENTIQQDPECHAIVKG-----VXTOQIOTQGQOLQVKETGVTKRTP 256
Db	187	-HKDAVNVERGASETESEMLEAFFFGEHEETKRLVAFQOEIHDHQPIKQR----- 237
QY	257	QKLFP---SPEIVKTHKAMER--LYAVFIDYEDKVSDEAVNKIRLDQEQLKEP 311
Db	238	---FVVERDEBLVKVKSLTBDKLKDVTL---FDQKQDNDLALK--EEVWGHFL 288
QY	312	PEADD-----YELESFNVVAKKEVRSIVIENEYKRCGDGDLTSLRNVCEVDMFKTLHG 365
Db	289	DEDBENETLVEKVVAILNDIKEVRLIADKEKIRDGRKDEIRPLESEVGLPRAH 348
QY	366	SALFRGQGOTQVLTCTMFDSLSEGKSDQVIAINGIKDKNFMHLHYRPPYATNEIGKVTG 425
Db	349	SGLFRGQTOQVLSVLTGAL---GDXQDILGIGPEVEKRFHMHNPNFSVGETGPVRA 404
XX	RESULT 13	
XX	ARU38175	
XX	DT	14-FEB-2002 (first entry)
XX	ID	AAU38175 standard; Protein; 721 AA.
XX	AC	AAU38175;
XX	PR	Salmonella typhi cellular proliferation protein #66.
XX	PR	Antisense; prokaryotic cellular proliferation protein; KW antibiotic; antibacterial; drug design.
XX	OS	Salmonella typhi.
XX	PN	WO200107955-A2.
XX	PD	27-SEP-2001.
XX	PR	21-MAR-2001; 2001WO-US09180.
XX	PR	21-MAR-2000; 2000US-191078P.
PR	PR	23-MAY-2000; 2000US-206848P.
PR	PR	26-MAY-2000; 2000US-20772P.
PR	PR	22-OCT-2000; 2000US-242578P.
PR	PR	27-NOV-2000; 2000US-233625P.
PR	PR	22-DEC-2000; 2000US-257931.
PR	PR	16-FEB-2001; 2001US-259308P.
XX	PA	(ELIT-) ELITRA PHARM INC.
XX	PA	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI	PI	Yamamoto RT, Xu HH;
XX	DR	WPI: 2001-611495/70.
XX	DR	N-PSDB; AAS56034.
XX	PT	New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids -
XX	PS	Example 3; seq ID No 13768; 51pp; English.
CC	CC	The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The anti-sense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from WIPO at
 CC format.wipo.int/pub/published_pct_sequences.

XX Sequence 721 AA;

Query Match 32.6%; Score 1158; DB 22; Length 721;

Best Local Similarity 39.9%; Pred. No. 9.8e-93; Indels 40; Gaps 11; Matches 270; Conservative 117; Mismatches 250; Index 18

QY 35 GSNRKLEISSGKLLARFADGSAVQSGDTAVMTAVSKTPKPSQ-FMPLVUDYRQKAAG 93
 22 GQHTXTKLETTGMMARQATAVMWSMDPAPFVTVWQGKAKGQDFFPLTVQYQERIYAG 81Db 94 RIPTNVLARRENGTSKDRKILTSRISDRSPRFLPAGYFYDPTQVLCNLIAVDCYNERPVLA 153
 82 RIGPSXKREKRPSEGETLARLIDRPRPFLPPEGFVNEVQVIATVWSVNQVNPDIVAM 141QY 154 NGASVAVLSDIPWNGPGHGVGAVRIGIDGKVYNTPKEMSTNLNVAGPKSQVMLE 213
 142 IGAASALSLSGPFPNGPGIAGVYINDQVYLNPTDQBLKSKLQDVLWAGT-BAAVLMVE 200

Db 214 ASEAENIQQDFCHAIKUGVKVTTQOIQIQLQVKECTVTKTPQKULFTPSBPKV---- 268

Db 201 SRAELSEDWTMIGAVVFGHEDQGQWVQAIIDLVKEAGKPRDWDQ----PEAVNDALNA 254

QY 269 YTHKLMERL---YAVTDYEDHKVSDNEANVKKRDLTDEQKKEKEPEADPYETTSFNV 325
 255 RVAALAESRSLSDAYRI---TDQBERIYAQDVKVSEETIEQLAEDETLIDANELGBILHA 309Db 326 VAKEVEVSIYVANEYKCDGRDITSLRNVSCVDMFKLHSALFORQOTVCTVFDL 385
 310 IEBKVNVRSLRVLAGEPRIDGRKDMRGLDVRGVLPRTHGSAFLPTKGETQALVATLGT- 368QY 386 ESGIKSDQVITAIANGIKOKNMLHYEPPPTAENIGKVTGLRREBHGHALAKALYPI 445
 369 ---ARDAQVDELMGERTDSLFHNMPPVSYGVTGNGSKGRREBHGHALAKRGFLAVM 425Db 446 P--RDPPTTIRVTEVSELESNSSSMASACGGSALALMDSGVIVSSAVAGVAGLGLVTKDPE 503
 426 PDMDKEPYTVRIVVSETESSNSSSMASVCGASLALMDAGVPIKAAYAGIANGLVLKEGD-- 483QY 504 KGEIEDYRFLDIDGLEDYDNGMDPKIAGTNGKIGITALQDIFLKGPIKIVMEAQASV 563
 484 ---NYVVLSDILGDBDHLGMDMDFKVAGSRUDGSLQMDPKIEGITKEIMQALNOAKG 538
 564 AKKEIIQIMNNTISKRSRKEENGPNVETVQPLSKRAKFPVGGGNLKKQAEQGTVIS 623
 539 ARLHIGVMBQINAARGDLSFAPRHTIKISTDKKDVIGKGGSVIRALTEEGTIE 598
 520 QDEETTSVFLAPTPSYHHEADPITSEICKDDQDQLEFGAVYATTEIRDTGVMKLYP 683
 599 IEDDGTVKIAATGEKAKYARRIEI---TAEIEVGRINKVTRIVGFAVAGG 653
 Db 684 NMTAVLAHNTQDLDNERL 700
 654 GKEG-LVHISQIADKRV 669RESULT 14
 AAY03792
 ID AAY03792 standard; Protein; 1034 AA.
 AC AAY03792;
 XX 11-JUN-1999 (first entry)DE S. aureus polypeptide.
 XX

Staphylococcus aureus polypeptide; thyroiditis; infective carditis; lung abscess; secretory diarrhoea; cerebral abscess; conjunctivitis; toxic shock syndrome; folliculitis; septic arthritis; antibacterial; H. pylori infection; gastric ulcer; adenocarcinoma.

XX Staphylococcus aureus.
 OS

DR EP905243-12.

DN 31-MAR-1999.
 XXPP 03-AUG-1998; 98EP-0306185.
 XXPR 05-AUG-1997; 97US-0055387.
 XX(SMIK) SMITHKLINE BEECHAM CORP.
 PABurnham MFR, Lonetto MA, Warren PV;
 XX

DR WPI; 1999-192867/17.

DR-N-PSB; AAX31862.

XX New essential polypeptides from Staphylococcus aureus useful for PT treating diseases such as infective endocarditis and toxic shock PT syndrome

XX Claim 31; Page 50-55; 70pp; English.

The invention provides new Staphylococcus aureus polypeptides (AAY03781-94) and the genes (AAX1851-864) encoding them. Host cells containing vectors comprising the nucleic acid sequence are used for the recombinant expression of the proteins. The polypeptides can be used to screen for modulators for use in antibacterial therapy. The polypeptides, their antagonists and agonists are used to prevent or treat diseases caused by S. aureus such as thyroiditis, lung abscesses, infective carditis, secretory diarrhea, cerebral abscesses, conjunctivitis, toxic shock syndrome, folliculitis and septic arthritis. Screening for the presence of the polypeptides may be used to diagnose, predict the susceptibility to, or stage the progress of these S. aureus diseases and diseases caused by Helicobacter pylori such as gastric ulcers and gastric adenocarcinoma. There is not much information known about the essential genes expressed by S. aureus during infection but these new polypeptides have been identified as essential. They can therefore be used to develop antibacterial compounds specific for those essential genes and this ensures the effectiveness of the compounds in killing S. aureus. In addition, these polypeptides can be used to effectively diagnose and treat infections and diseases caused by S. aureus without the risk of development of antibiotic resistance. The present sequence represents a S. aureus polypeptide which has homology to a polyribonucleotide nucleotidyltransferase.

SQ Sequence 1034 AA;

Query Match 32.5%; Score 1156; DB 20; Length 1034;
 Best Local Similarity 39.3%; Pred. No. 2.7e-92; Indels 58; Gaps 11; Matches 265; Conservative 113; Mismatches 239; Index 18QY 24 SAGSRAVAVDGLGNRKLEISSGKLLARFADGSAVQSGDTAVMTAVSKTPKPSQFMPFLV 83
 2 SOBKVKTPEWAGRSLTISGQALQANGAVLVRVYGTFLVUSTASKPBRGDGFPLTV 61Db 84 DYROGAAAGRPIVLLRREVGTSKELTSRISDRSPRFLPAGYFYDPTQVLCNLIAVD 143
 62 NYEEKMYAAGKIPGKPKBGRPGDADTARLDRPRFLPFGYKQHVOQIMMNLSD 121QY 144 GVNEPDVLAINGASVALLSLSDIPWNGPGHGVGAVRIGTIDGEGVYVNPRKEMSSSTNLVAG 203
 Db 122 PDCSPQWAMIGSWSVSDIPFQGFLAGVNVGIDGKVIINPTVEKEVRSRDLDEVAG 161
 QY 204 APKSQIVMLLEASAENTLQQDFCHA1KVG-----VKTQHQIQQIQLQVKTETGVTKTP 256

XX PT
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
PS
Example 3: Seq ID No 10312; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC
CC
ftp://wipo.int/pub/published_pct_sequences.

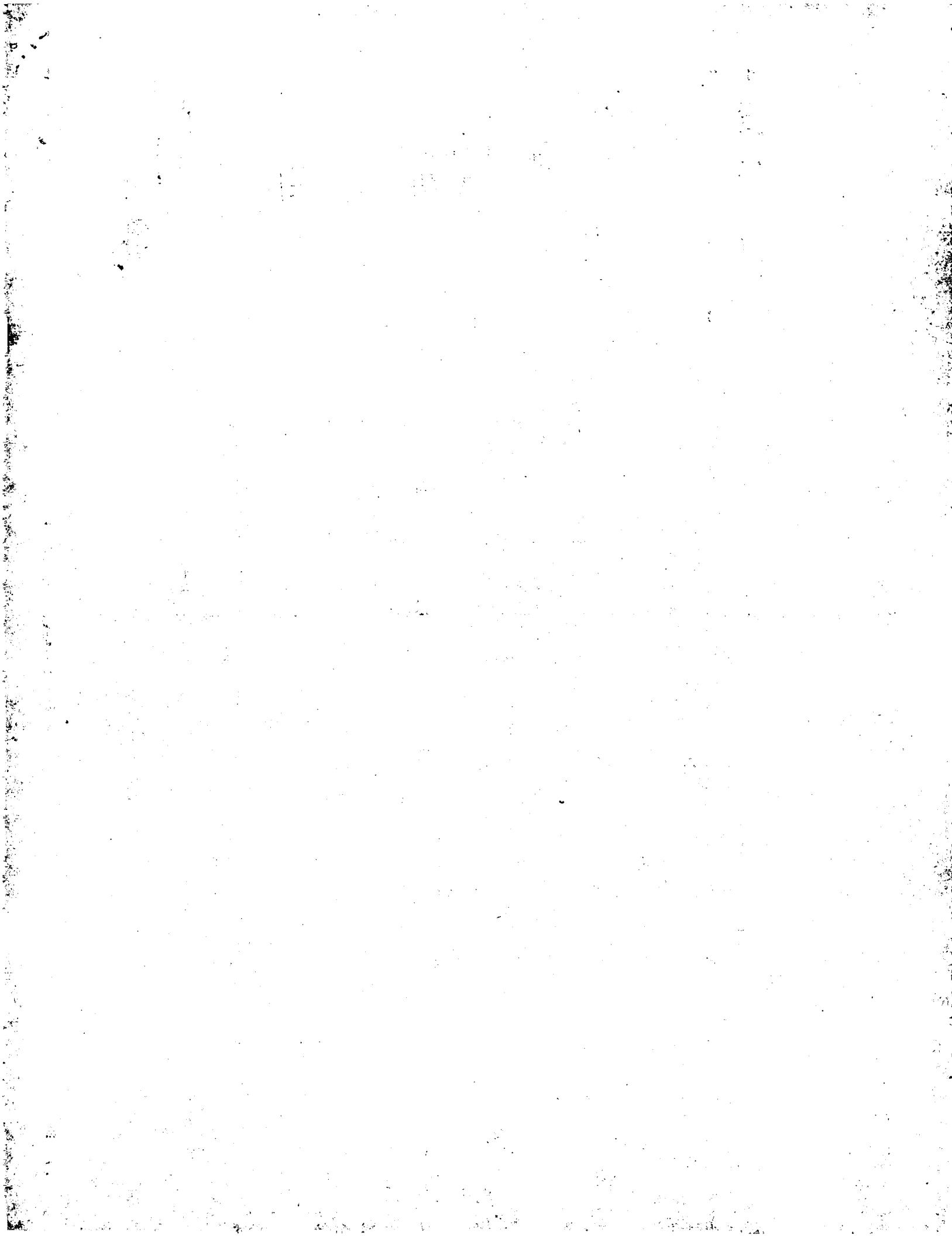
Thu Jan 8 13:54:16 2004

us-09-907-907a-42.rag

Page 13

Db 614 DDGTVKIAATGEGAKHAIRIEI-----TAEI E VGRVVTGKVTRIVDFGAFVAIGGK 668
Qy 686 TAVLIAINTQOLDNERL 700
|: : | : : |:
Db 669 EG-LVNTISQIADKRV 682

Search completed: January 8, 2004, 10:38:47
Job time : 51 secs



QY 366 SALFORQOTQVLCITWFDSSLESQIKSDQVITAINGIKDKNFMHYEPPTATNEIGKVG 425
 ; US-09-107-532A-5868
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5868
 ;
 Db 349 SGLETFRQTOALSVLITGLA-...-GDYQLIDGLPVEPRKFMHNPFNSVGSTGPVRA 404
 ;
 QY 426 INRRELGHGALAEKALYVLP-...-RDPFTTRVTSEVLESNGSSMASACGGSLALMDGV 483
 ;
 Db 405 PGRRIGHGALEKALYVLP-...-RDPFTTRVTSEVLESNGSSMASACGGSLALMDGV 464
 ;
 QY 484 PISAVAGVAVLGLVLTPTDPERGEEDYRLITDLEDYNGMDPKIAGTNGKGTALQAD 543
 ;
 Db 465 PIKAVAGVAGLVLVTRDD-...-SYTILDICQMEDALGMDPKVAGKDGITAQMD 517
 ;
 QY 544 IKLCPIPKIVMEAQASAKKELQIMKNTISPKRASKRKENGPVWETVQVPLSKRAF 603
 ;
 Db 518 IKDQSLTREVTEALEBOARQRLAIDMDMLHITBOPRELSAYAPKVUTSINPDKIRD 577
 ;
 QY 604 VGGGGYNLKKQAEQGVTISQVDETFSVFAPTSVMHEARDFTETCKDDQBOQLEFGA 663
 ;
 Db 578 IGGGKKEKINELIDEGVQDIEQDTEIPGAVDQMINRAKKEIDTR-...-EAVCQ 632
 ;
 QY 664 VYATATTEIRUTGVWVLYRNMTAVLHNTQDNERLN 701
 ;
 Db 633 VYHAKVRIEKVGFVBLFFPKDA-LHHSQISQERIN 669
 ;
 RESULT 2
 US-09-107-532A-5868
 ; Sequence 5868, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ;
 APPLICANT: Lynn A. Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 ;
 CORRESPONDENCE ADDRESS:
 ;
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02454
 ;
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 ;
 CURRENT APPLICATION DATA:
 ;
 APPLICATION NUMBER: US/09/107.532A
 FILING DATE: 30-Jun-1998
 ;
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ;
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 ;
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5507
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 5868:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 750 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...750
 ;
 RESULT 3
 US-09-328-352-8112
 ; Sequence 8112, Application US/09328352
 ; Patent No. 656258
 ; GENERAL INFORMATION:
 ;
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ;
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTW99-03PA
 CURRENT APPLICATION NUMBER: US/09/328.352
 ;
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 8112
 LENGTH: 705
 ;
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 ;
 US-09-328-352-8112

Query Match 29.9%; Score 1065; DB 4; Length 705;

Best Local Similarity 36.8%; Pred. No. 5.5e-91; Mismatches 250; Conservative 124; Indels 44; Gaps 14;

Qy 35 GNRKLETTSGKLLARFADGSAAVWQSGDTAAMVTAVKPSQSPMFLVNUYDROKAAAGR 94

Db 22 GHOQVVLBTGRVARQANTVLTGGVTVLAVVAAAPPKAGODPFPPLTVNQEKQVAAAGR 81

Qy 95 IPIVNLREVGNSDEKILTSRDRSTRPLPAGYFDTQULCNLLAVDGNPENPDVLA 154

Db 82 IIGGGYKREGRASEBAETLISRSLDRPRLPFPGRGYNMIIQVATWSSDKTNEADIA 141

Qy 155 GASVALSLSIDIPNGPGCAVRIGIDGIVVNPTRKEMNSSTLNLUVAGAPKSIQVMLEA 214

Db 142 GTSAAALAIAGTTPRGPIGAARVGLINGEYVLUVPFQMAQSPDLVWAGT-ESAVLIVES 200

Qy 215 SAENILOODFCHAIAKVGVYKTOORIQGQLYKETSYTKRTOKLETPSPBIVKTHLA 274

Db 201 BAKELSERDQMLGAVLFQHDEMOMIAQIAINEFAAAGA---KPSDWAPA-----HN- 248

Qy 275 MERLYAFTDYEVDKVS- -DEAVNKIR- -LDT- -BQEKEKFPEDPYLETESTVVA 327

Db 249 EBLRALKUKEAFAEKCFBAYTIAVKOYRAALDAHLAEAVQFUPFEDVSDIADETYLP 307

Qy 328 KEV-FRST- -VLANEKYKCDGRLTSRANSCEVDMKTLQISALFORQOQLCWTED 383

Db 308 EDLKTYRTWTRDNLISGKPRIDGRDTKTRALDQVQGULERANGSALFTRGEQALVTLG 367

Qy 384 SLSGQIKSDQVATANGIKDQKFLMHPAYTNPBKGKVTLNRLBIGHQSLAERKLYP 443

Db 368 NTRDAL---WMDTLAGTKTNFMHLHNFPAVSGETPRESQPKRKRIGHQSLARRGVQA 423

Qy 444 VLP- -DFPFTIRTVTSELESNGSSMASACGSISLAMDSPVPISSVAGVIAGLYTKD 501

Db 424 VLPADADKPVYVIRIVSITDSESNGSSMASVGASLISLMDSPVPLKAVAGTAMGLK- - 480

Qy 502 PEKGEIDYRILTDIQLGIEDTNGMDMFRKIAITNGKITALQDQIKLPGIPKIVMEAIQQA 561

Db 481 ---EGERFAVLSIDQEDHGDMPKVVASSANGITALQDQIKIGITBIMEVALNQA 536

Qy 562 SWAKKBLIQMKNTISKPRASRKENGPNVETVQVPLSKRAKEVPGCGYNLKKLQATGVT 621

Db 537 FAGRMMHILNEMKVKISPARPEISMAHTPEVITINPDKIRDVIGKGGATIRQITEKAA 596

Qy 622 ISQVDEETFSVAPTPSMHARDFTIECKDQEQOLEFGAVYATTITERDGTGVMKL 681

Db 597 IDEDNGTVRVIGETKAAKAAJAKTOI- -TAEVERGKTYDGVIRVIEFGAVN 651

Qy 682 YPMNTAVLHLNTOLDNERL 700

Db 652 MPG-TDGLHLHSQISNERI 659

ORGANISM: *Pseudomonas aeruginosa*

US-09-252-991A-27283

Query Match 27.5%; Score 976.5; DB 4; Length 697;

Best Local Similarity 35.6%; Pred. No. 1.1e-82; Mismatches 230; Conservative 130; Indels 53; Gaps 15;

Qy 86 ROKAANAGRIPNTYLR-----EVGTSDEI-----LTSR1RDRSTRPLPAGYFY 131

Db 45 RPKAATSSRCFCCTRKPTPOAVSPAVSSVQDVRRENTSRSLDRPRLPFPFGPM 104

Qy 132 DTQVQNLATDGVNPERDVLINGASVALSLSIDIPNGPGCAVRIGI-IGEYVNPTRK 190

Db 105 EVQVVCVTVVSTNKEDDPIAMIGTSALATSGIPFAGPPIGAARVGFHPEGYLNPPTYE 164

Qy 191 EMSSTSTLNLUVAGAPKSIQVMLEAASBNILQODFCHAIAKVGVYIQQICIQOLVKE 250

Db 165 QLOSSLDMVWAGT- EDVAVLNVESADELTDOMQFLAHDEFOAVIRAKLAAEAG 223

Qy 251 VTK- -RTPQKLFPSPEVIVKTHKLA MERLYAFTDYEVDKVS- -RDEAVNKIRLD 302

Db 224 KPAWDWKAPAB-NTVLUWATKELGBlASQYTI- -TIKODRINRIGLSDQAVLFA 280

Qy 303 TEEOLKEKEPFPADPYLETSPNWKAVEVFISIVLBYKRCGDRJTSRANSCEVDMFKT 362

Db 281 EB- -GKFPPAS- -EVKDVFGLLBYRTVRENIVNGKPRIDGRDTARTVPLRIVGVLGK 333

Qy 363 LHGSALFQRGOTQVLCVTDTSLESGKSIQVITANGIKDKNFLHNPYPPATNEIGK 422

Db 334 THGSALFTRSTQALWATG- -ARDQDILTLEGERDAFLMLHYNPPSYGECGR 389

Qy 423 VTGLNRELGIGLAAEKKALYVPIR- -DFPFTIRTVTSELESNGSSMASACGSISLAMD 480

Db 390 MGSPGRERIGGRLLARRGVAAMLPQDDEPFYTRVSEITESNGSSMASVGASLISLMD 449

Qy 481 SGVPTISSAVGVAIGVTKDPEKBEIDRFLTIDGIDYNGMDKLAGTNGKITAL 540

Db 450 AGVPKVPAVGIAJGMVK-----EGKFAVLTQDGDHLDGMDFKVAGTDGVT 502

Qy 541 QADIKLPGIPKIVMEAIQQAQSVAKKELQIOMKNTISKPRASRKENGPNVETVQVPLSKR 600

Db 503 QMDKINGTVERBILALGQALEARLNLLGOMVNQIAKPAELSAINAQTLQMKUDKI 562

Qy 563 RDVIGKGGATIRGICBETKASIDIEBDGSVKGIVGETKEAAKURVLA- -TAEAE 617

Db 611 FGAVTTATBIRDGVWVLYPNTA VLAHTOLDNERL 701

Qy 618 IGKIVVGKVERIVDFGAFVNILPGKDG-LVHISQISDKR 657

RESULT 4

US-09-252-991A-27283

Sequence 27283, Application US/09252991A

; Parent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196_136

; CURRENT APPLICATION NUMBER: US/09/252, 991A

; CURRENT FILING DATE: 1999-02-18

; PRIORITY APPLICATION NUMBER: US 60/074, 788

; PRIORITY FILING DATE: 1998-02-18

; PRIORITY APPLICATION NUMBER: US 60/094, 190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27283

; LENGTH: 697

; TYPE: PRT

; OTHER INFORMATION: xaa=unknown or other

RESULT 5

US-09-198-452A-1073

Sequence 1073, Application US/09198452A

; General Information:

; APPLICANT: Giffels, R.

; TITLE OF INVENTION: *Chlamydia pneumoniae* genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: and treatment of infection

; FILE REFERENCE: 97110-003-599

; CURRENT APPLICATION NUMBER: US/09/198, 452A

; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 1073

; LENGTH: 568

; TYPE: PRT

; ORGANISM: *Chlamydia pneumoniae*

; FEATURE:

; NAME/KEY: SITE

; LOCATION: 1..568

Best Local Similarity 21.1%; Pred. No. 0.0014; Matches 135; Conservative 79; Mismatches 236; Indels 189; Gaps 31;

QY 374 TQVLCTWTFDSLESGKSDQVITAINGIKDNFMFLHEFPPYATNGKVTGLNRELGH 433
 Db 677 ---ADVLFDYD--VSNDQDSIYPOYDRC--QFDKEMTPNSPG---- 715

QY 434 GALAEKALYKPYPRD----FPP----TIRVSEVLENSG 465
 Db 716 ---YPTITFDENTNSTYDFGKTNKRYIEVGANGNWIDVPTLYITGTAKEPQSN 767

QY 466 SSMASACGGSLALMDSCGPSSA--VAGVAGLVLTKDPEKGIED--YRLTDLIGIED 521
 Db 768 NNEGSA--SVSVQNEALDISATOAANPTLKVNKTIVTQNIDNKTHRKNPTIEL-- 822

QY 522 YNGDMDFKLAGTNKGITALQD--IKLPGIP--IKVMEAIQQASVAKKEILQIMKTI 576
 Db 823 ---TPKGTTNAQDIDNSITVKGVPEADASLEKTINGAKVIFKQYDLYTENITI 871

QY 577 SKPRASKRKGKGPVWETV----QVPLSKRAKFVGGYGNLKKLQAEVTIS---- 623
 Db 872 EYNTVSAANAGQYTTIDSETLNQMSASKKVTAPITLKFSBGDAEGIVYLATFTY 931

QY 624 -QVDEBETFSV--FAPTPSMWHEARDPTEICKDQDQOQLELFGAV---YTATTEIRD 674
 Db 932 HNVEDENQAIKVSFELIDNWTATEFTT----DEKGQSFDAIMTGDYTLRVNVPQ 986

QY 675 -----TGWVNLYY--PNMTAVLHNTQDNERLNTI 702
 Db 987 BYSVDBEYLTKAIKLVKGDNQKPLTKT-IDHSRLQV 1024

RESULT 8
 US-09-071-035-402
 ; Sequence 402, Application US/09071035
 ; Patent No. 6448043
 ; GENERAL INFORMATION:
 ; APPLICANT: Gil H. Choi
 ; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
 ; NUMBER OF SEQUENCES: 496
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/071, 035
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: A. Anders Brookes
 ; REGISTRATION NUMBER: 36, 373
 ; REFERENCE/DOCKET NUMBER: PR369P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 402:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1448 amino acids
 ; TYPR: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-071-035-402

RESULT 9
 US-09-252-991A-23776
 ; Sequence 23776, Application US/09252991A
 ; Patent No. 6551755
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074, 788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094, 190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 23776
 ; LENGTH: 252
 ; Query Match 3.9%; Score 140; DB 4; Length 252;
 ; Best Local Similarity 22.1%; Pred. No. 7.7e-05;
 ; Matches 62; Conservative 48; Mismatches 109; Indels 62; Gaps 11;

Query Match 3.9%; Score 140.5; DB 4; Length 1448;
 ; Query Match 3.9%; Score 140.5; DB 4; Length 1448;
 ; Query Match 3.9%; Score 140.5; DB 4; Length 1448;

QY 331 FRSTIVLNAYK----RCGDRLTLSRNVSCEVDMFKTLHGSAFLQRGOTQVLTCTVTFDSL 385

Db 1 WRPPIFLSQRKRISTMRNPKPSGRAODQPLRPIRTRHVKHAGSSVLFEGDTKVICV---SA 57

QY 386 ESGIUSQDVITAINGIKDKHFMHLRFEPPVATNEIGKVYGLNRLBIGHALAEKAL--YF 443

Db 58 ESGVP----RFLKGQGQGMILTAEVGMLPRSTGE----RNQREASRGRQGGRTLEIOR 106

QY 444 VPIRPF-----PFTTRVTSVLESNSGSSMASACGGSLALMDS-----G 482

Db 107 LIGRLRAALDLISLKGENTYVDPCTVQADGGTRASITGATVALLDALAVLKRGALKG 166

QY 483 VPISSAVAGVIALGLVTKTDPEKGETEYDYLTDIGIEDYNGDMDFKIACTNGITALOA 542

Db 167 NPLKQMVAAVSVGIVYQGV-----VLDLQYLEDSEAETDLNVMTAG-GFIEV 214

QY 543 DIKLRGIP----IKVMEIAIQNSVAKKEILQIMNKTK 578

Db 215 QGTAEGAPFRPRAELNAMLELAAQ--GMOBLFELORAALE 252

RESULT 10

US-09-252-991A-27186

; Sequence 27186, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenstein et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196_136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: -US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO: 27186

; LENGTH: 191

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-27186

Query Match 3.9%; Score 137; DB 4; Length 191;

Matches 33; Conservative 18; Mismatches 38; Indels 4; Gaps 2;

QY 35 GNRKUBISCKLARFADGSAWVOSD-TAVMTAVSKTKPSO-FMPLVYDROKAAG 93

Db 94 GQSTVTLERGTRIARQQTGAVINLVTMDVSVLTVVAGAKSPAEGRDFEPLSVHYQEKTYAAG 153

QY 94 RIPTVYLRREVGTSDK--ELTSRIDIISRIP 123

Db 154 RIPEGFFKRGGRPSERKGNPDAPDRPSDPSAVP 186

RESULT 11

US-09-328-352-5628

; Sequence 5628, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: ACINETOBACTER BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO: 5628

; LENGTH: 244

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5628

Query Match 3.6%; Score 128.5; DB 4; Length 244;

; Best Local Similarity 21.2%; Pred. No. 0.0008; Mismatches 109; Indels 75; Gaps 9; Matches 61; Conservative 43; Mismatches 109; Indels 75; Gaps 9;

QY 336 LNEYKRCGDRDTSIRNNSCEVDMFKTLHGSAFLQRGOTQVLTCTVTFDSLQSVI 395

Db 3 LGRIMRINDORALDOLPREYKTRNTRYTRYAGSVLVEFGHTKVLCTASID-----NSVP 54

QY 396 TAINGIKOKNFMLYEFPYATW-----IGVTKGN--RRELHGALAKALYVVI 445

Db 55 RFLKGQGQGWVTEYGMPLRSTSICRDCBAECKQTKTQETORLIGSLRANWDLKLG 114

QY 446 PRDFPFTTRVTSVLESNSGSSMASACCGSLALMDS-----GVPSSAVAGVAT 494

Db 115 EN---TITIDCDVIQALOGIRTASITGAAVALVDAMNLLAOKKIKDQPLKGLVAAISV 170

QY 495 GLVTKTDPEKGETEYDYLTDIGIEDYNGDMDFKIACTNGITALQDIDKUGPIKIV 554

Db 171 GM-----YD-EVILDCYEDSNDQTDLNWVMOAG----- 201

QY 555 MEATQOASVAKKEILQIMNKTKSPRASRKEPENGPVETVQVPLSKRAK 602

Db 202 -----ERIEIQCATAEDKP-FTRAQSNAMEKGAELIK 236

RESULT 12

US-09-134-001C-3979

; Sequence 3979, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-03-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO: 3979

; LENGTH: 871

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

; US-09-134-001C-3979

Query Match 3.6%; Score 127.5; DB 4; length 871;

; Best Local Similarity 20.0%; Pred. No. 0.0099;

; Matches 145; Conservative 116; Mismatches 276; Indels 189; Gaps 33;

QY 10 DRAITQOLQYRALMSSAGER-----AVADLGNRKEIISSSKLARTADG-SAVVQSGD 60

Db 235 DKTIEFLDLSALVAGAKRGEFEERLKV-----KEVKESEGRILFIDEHMLVAGK 289

QY 61 TAVMTAVSKTKPSO-----FMPLVYDROKAAGAAGRIPYLREBVGTSKELTS 114

Db 290 TDGAMDAKMLKMLKARLHCIGATINNEYREKESALERPF--OKVGSEPDVENT 347

QY 115 RIDRSIRPLFPAGYFPTQVLNCNLAVDGWNEPPDVLALINGASVALSLSD----IPW 167

Db 348 ISTIRGLKERYEVH-----GVRQDRAV-----AAAEELSDRYITDREFLD 389

QY 168 NGPYGAVRTGID-----GEYVNPT-----RKEMSSSTLNUVAGAPSKOIV 210

Db 390 KA-----IDLVPOACATIRTEMGSNPFLDQNRVWMOLETEESALKNESDNASKHLR 443

QY 211 MLEASANTLQOPCHAKVGVKTYQIIGOLVKEGTGVTRPQKLFPSPEIVCYT 270

Db 444 ELOPELSN--EKEQSSIKSRVQKBEKIAVQKVRKALDSSRQALEAQTEG----N 495

271 HKLAMERLIAVETDYEHDKVSRSRDAV-NKIRLDPEFEQLEKQKFPEADPYEIESNIVANG 329

495 LEKAKBLOQYGTIPOLEKELEQBFEEFQDGEDGSDERMRVVSDEB1GDIVS0W----- 550

Db 330 VFRSIVLNHYKRCGDRLLSLRNNSCEVDMFKLIGSALFORQGOTQVLCTVFDSSLESGI 389

Db 551 --GIPVSKVLETEREKLUSLSDI-----LHKRKV--GDKAVDLV----- 586

Db 390 KSDPOVITAINGIKDRKFMHLHYEPFPYATNEIGKV---TGLNRERLGHGALAEKALYLV 444

Db 587 -SDAVVRARPGIKOPN-----RPIGSFLPLGPTGVGKTEL-----KSLASS 627

Qy 445 IPDFPFITRV-TSEVLENGSSMASACCGSLAMDSGVPISSAVAGVAIGLYTKPE 503

Db 628 LFDSKHKMIRIDMSYMEKHAWSRNLIGAPPYVG-HD8GCOLTERAVRNPNVSYVILDE 686

Qy 504 KGETEDYRUDTDLG---TEDYNG-DMDFK---IAGTNKGITALQADTKLPG---IPI 551

Db 687 KAHSDVFVNLLQIDDEGRITDKSKRSVDFKNTIIMTSWISQVLLNKDAGISDDE 746

Qy 552 KIVNEHAIOASVAKKEILOIMNKUTSKPRASKRKGNGPVPVETVQVPLSKR----- 600

Db 747 KAVMDSLH---AYFKEPEILNEMDDVLFKPLSVDDMSMTVKILQNLQNLDDOHISIEVT 804

Qy 601 --AKFVGCGIN-----LKKI---QATGV-----TISQD---BETP 630

Db 805 EKACKWLGSEAYEPOFGARPLKRVQROETPIARMMIKESLPEGGTIKVDLNDNKELP 864

Qy 631 SVFAPT 636

Db 865 KVKVKT 870

RESULT 13

US-09-107-532A-5705

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 893-5007

TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 5705:

SEQUENCE CHARACTERISTICS:

LENGTH: 471 amino acids

TYPE: amino acid

RESULT 14

US-08-999-774A-13

Sequence 13, Application US/08999774A

Patent No. 6274312

GENERAL INFORMATION:

APPLICANT: Gish, Kurt C.

APPLICANT: Seghezzi, Wolfgang

APPLICANT: Shanahan, Frances

APPLICANT: Lees, Emma M.

APPLICANT: McClanahan, Terrill K.

TITLE OF INVENTION: Intracellular Regulatory Molecules; Related Reagents

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute

STREET: 901 California Avenue

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent-in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/999-774A

FILING DATE: 10-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/032,818

FILING DATE: 11-DEC-1996

ATTORNEY/AGENT INFORMATION:

LOCATION: (B) LOCATION 1...471

SEQUENCE DESCRIPTION: SEQ ID NO: 5705:

US-09-107-532A-5705

Query Match

Best Local Similarity 3.5%; Score 124.5; DB 4; Length 471;

Matches 66; Conservative 56; Mismatches 106; Indels 85; Gaps 15;

Qy 37 RKEIISGKULARFADGSAVQVSGDTAVMTAVSKTKPPSQFMPVLYDVRQKAAAGRIP 96

Db 81 AYSMLPRATSTRNRESSKGKLFTGTEIQRLIGRSIRAVDLEKLGERSIVDCDVIQ 140

Db 32 RKLQLQ-TVFKYFEGSVWIRFGDTIVCSATINSVPP--FL-----RETGGMWT 80

Qy 142 VDGYNEDVLAINGASVAL-----SLSDIPWNGPGAVRIGT-DGEYVUNPTR 189

Db 141 ADGTR--TASITGAFVAKLAKTSLRSRPIKSHAAVSVGLIPDGTCVTDY 198

Qy 190 KEMSSSTLIVVAGAPKSQLVMLEASAENILQQDFCHAIVKGVKTQOIQG--IQOLV 246

Db 199 QEDSAALVDMNLVWMTESGKFEIOTGE-----BAYFDEQLENLIFFGKNAEIDL 250

Qy 247 KETGVTKRTPQKLFPRSPBVKYTHKLAMERLYAVFTDEHDKVSRSDEBAKTKRIDLTE- 304

Db 251 KE-----OK----- -HALUTERFAQND-BRIETKTIIAATRNP 281

Qy 305 --EQLKERPEA 314

Db 282 GKAEEFRNMFKEA 294

NAME: Ching, Edwin P.
 REGISTRATION NUMBER: 34,090
 REFERENCE/DOCKET NUMBER: DX0645
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 486-1200
 INFORMATION FOR SEQ ID NO: 13:
 LENGTH: 1306 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-999-774A-13

Query Match 3.5%; Score 124.5; DB 3; Length 1306;

Best Local Similarity 19.7%; Pred. No: 0.038; Mismatches 149; Conservative 112; MisMatches 239; Indels 257; Gaps 38;

Qy 73 PSPSQFMPFLVVDYRQKAAAGRQPTNLYRREVGTSKE--ITSRIDR----SIRPL 124
 Db 530 PLFNGQTLPLNS-----PT-----LDQERHILVLRVQIDRILVKKLDDLVRP- 569

Qy 125 PPGAGFYDFTQVLNCNLIAVDGNEPDVLAINGASAVASLSLSDTPWNGPFGVAVRIGITDG--E 182

Db 570 -YVHKILVIEPLID--EDYARVEGLEISNLAKAGLAMTISWRPDDINMDE 622

Qy 183 YVWNPNTRKEMNSSTANLWVA--GAPKSQTMLESAENTIQQDCEHAIKVGKVQTOQ-- 237
 Db 623 YVRENTTARATA-----WVASALGIP-SLIPFLKAVCKS--KKSQWQARHGTGIVQIA 672

Qy 238 -IIGQIQLQVKEGTGVTKRTQKLFPSPEVKYTHKLAMERLYAVFTDYEHDKVS 291
 Db 673 ILMGCAILPHLRSLV-----EII-----EHGLVD 696

Qy 292 RDEAVNKIRLDEEQKEKIPPEADPVEIIFSSFNVVAK----- 328

Db 697 EQQKURTISALAIANLAE--AATYG-IESFDISLKLPLWKGIRQHRGKLAELKAIGY 752

Qy 329 -----EYPERSTVLINEYKRCDCDRLTSLRNVSCEV-- 357

Db 753 LIPLMDAELYANYYTREVMILIREQSPDDEMKCILVKVVKQCCGTDGVANVYKTELP 812

Qy 358 DMFKKL--HGSSALFORQQTQVLTCTVTPDSLESGIKSDQVITAI NGIKDNFMLYEFP 414
 Db 813 PFPKHFQWHRMAILDRNRYRQLUDT'VELANKVQHAEIISRIVDLDK--ABQYRK 868

Qy 415 YATNEIGKVIG--INRRELIGIGALAKALYPVIPRDPFPFTIRVTSEVLESGSSMASA-- 471

Db 869 MVMETIEKINGNLGAIDH-KLESQDIL--YAFQEQTTERSVMANGFGTVNALG 924

Qy 472 -CGGSLSALMDSGVPTISAVAGVAGLVTKDPEKGIEDYDYLTDI-LGIE 520

Db 925 KRVKPVLPQICGTWLRNKK--SAKVRQGADLJSRTAVVMKTCQEEKLMGHGVVLY 981

Qy 521 DNGDMDPKIAGTNGKITAQDILPGI---PIKIVM-----EATQO--- 560
 Db 982 EYLGSEYPEPEVGISLG--ALKAIKVIGMHNTPPIKDLRFLTPIKRNHRKVQENCID 1039

Qy 561 -----ASVAKKEILQTMNKIKTSKPRASRKE--NGPVWETVQVPLSKRAKFVGGY 609

Db 1040 LUGRTRADRGAAEYVSAREWMRICKFELLELLKKHKAIRATWNT---FGYIKAIGPHDV 1095

Qy 610 -----NLKQOAE--TGVTISQDEETTSVFAFTPSVHEAR----- 644

Db 1096 LATLNLNLKQOBERQNCVTTAIAV-AETCSPFTLPAJMNTRVPELAVQONGULKLS 1154

Qy 645 --DPTEICKDDQEQQLEFGAVVATITRDRGWM 678

Db 1155 FLFVEYIGEMGD-----YIYAVTPLLEDALM 1180

US-08-769-309A-5

Sequence 5, Application US/08769309A

Patent No. 5741890

GENERAL INFORMATION:

APPLICANT: Scott, John D.

APPLICANT: Nauert, Brian J.

APPLICANT: Klauck, Theresa M.

TITLE OF INVENTION: Protein Binding Domains of Gravin

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower/233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

COMPILER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patient in Release #1.0, Version #1.3.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/769,309A

FILING DATE:

CLASIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5741890and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/33451

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 25-3856

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1780 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-769-309A-5

Query Match 3.5%; Score 116; DB 1; Length 1780;

Best Local Similarity 19.6%; Pred. No. 0.41; Mismatches 269; Indels 182; Gaps 35;

Matches 134; Conservative 99; MisMatches 269; Indels 182; Gaps 35;

Qy 40 EISSSKLAFLPADGSAAVSGDGTAVWVATSKTPSPSPDFMLVYRQKAAAGRQPTN 99

Db 882 BLSQSQVHMA--AAVADGTRAAITI--EERSPSNISASVY-- 923

Qy 100 LRRVGTSKDEILTSRIDSIRPLPAGFYDPROQVNLAVDG--VNEPDVL--IN 154

Db 924 VEASALLTTEEVLEREVAEERBPPV-----TEPLPENREARGDTVSEABLTPEAVT 976

Qy 155 GASVALSLSIDPWNPGPVAVRIGIDGKVVNPRKEMSSSTLNVLVAGPKSIVMSEA 214

Db 977 AAETA-----GPIGS-----EEGTAEAAETTEMV-----SAVSQJTD 1010

Qy 215 SAENIQLQDPRCHAIKV-----KTYQIQIQQIQLQVKGTVKRT--PQKLFTPPE 265

Db 1011 SPDTBEATPVQEVEGGVFDIEOERRQEVQLAQAAKEKEESOLPGTGPEDVLQP-- 1067

Qy 266 IVKVTTHKLAMERLYAVFTDYEHDKVSRAEVNKKRILDPTEBEOLKEKFPADPVETI-- 320

Db 1068 VORABAERPEBQEAZASGLKKEKTDV-----LKVDQAEAKTEPFTQG--KVVGQTP 1116

Qy 321 RSVNVAKEVFRSVLINEYKRCDCDRLTSLRNVSCEVMKTLHGSALFQGQTQULCV 380

Db 1117 ESB-KAPQVTESTESSELV-----TTQCAAS--TTLAG----VKSQBMW---- 1152

Qy 381 TFDSELSGKTSQDQVITATNGIKDNFMFLYEPYVATNBIGKVGLNRELGH----GA 435

Db 1153 --MEOQAPPDSETPTPSETDGTSPV-ADFDAPGTTQKDELEVEIHENEVHLVPGT 1207

QY 436 LAB---KALVAVIPDPFPTRVTSBVLENGSSMASAGGSILALMDSCVPIASAVG 491
 Db 1208 BAAVAAQKERPPAPSSFVFO-----ETREQSKMED---TJERTDKEVSET--- 1252
 QY 492 WAIGLVTKTDBKGEBDY--RLTILIGISDYGMDPKAGTNKGITALQADIKLPGI 549
 Db 1253 --VSILSKTEGPO-EADYOABRKTDVPPFGLEGSDITGTVSREKV-----EVALKG- 1304
 QY 550 PIKIVMAIQQASVAKKEILOQIMNKTSKPPRASKRENGPVVETVQVPLSKAKFVGPGY 609
 Db 1305 -----EGTEBAECKODALLQSHAKSPP-----SPVREMVQVEREKTEARPTHV 1351
 QY 610 NLKKGQAEETGTTISO-----V-----ERPSV-FAPTSVMMHARDFTEICK 652
 Db 1352 NEBKLEHETAVTVSEEVSKQLQTAVNPILIDAKEVSSLEGSSPPCLGQE-----BAVCT 1406
 QY 653 DDQEQQLEFGAVVATITBIRDG 676
 Db 1407 KIQVQDSEASSTILTAAEEEKVLG 1430

Search completed: January 8, 2004, 10:41:27
 Job time : 24 SECs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model
Run on: January 8, 2004, 10:38:52 ; Search time 37 Seconds
(without alignments)

Scoring table: BL0SUM62 Gapop 10.0 , Gapext 0.5
US-09-907-907A-42

perfect score: 3557

Sequence: 1DGFFLLPDRRALTQLQVRA.....TAVLLHNTQLDNERLNILP 705

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgna_6/prodata/1/pupaa/US07_PUBCOMB.pep:*

2: /cgna_6/prodata/1/pupaa/PCT1_NEW_PUB.pep:*

3: /cgna_6/prodata/1/pupaa/US06_PUBCOMB.pep:*

4: /cgna_6/prodata/1/pupaa/US05_PUBCOMB.pep:*

5: /cgna_6/prodata/1/pupaa/US07_NEW_PUB.pep:*

6: /cgna_6/prodata/1/pupaa/PCTUS_PUBCOMB.pep:*

7: /cgna_6/prodata/1/pupaa/US08_PUBCOMB.pep:*

8: /cgna_6/prodata/1/pupaa/US09_PUBCOMB.pep:*

9: /cgna_6/prodata/1/pupaa/US08_PUBCOMB.pep:*

10: /cgna_6/prodata/1/pupaa/US09_PUBCOMB.pep:*

11: /cgna_6/prodata/1/pupaa/US09C_PUBCOMB.pep:*

12: /cgna_6/prodata/1/pupaa/US09_NEW_PUB.pep:*

13: /cgna_6/prodata/1/pupaa/US10_PUBCOMB.pep:*

14: /cgna_6/prodata/1/pupaa/US10_PUBCOMB.pep:*

15: /cgna_6/prodata/1/pupaa/US10C_PUBCOMB.pep:*

16: /cgna_6/prodata/1/pupaa/US10_NEW_PUB.pep:*

17: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

18: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

19: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

20: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

21: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

22: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

23: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

24: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

25: /cgna_6/prodata/1/pupaa/US60_PUBCOMB.pep:*

PRIOR APPLICATION NUMBER: US 09-243,277
PRIOR FILING DATE: 1999-02-02
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 42
LENGTH: 705

SUMMARIES

Result No. Score Query Length DB ID Description

1 3557 100.0 705 11 US-09-907-907A-42 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

2 3557 100.0 705 11 US-09-907-907A-44 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

3 1191.5 33.5 696 9 US-09-815-242-5443 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

4 1191.5 33.5 698 9 US-09-815-242-1225 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

5 1158 32.6 721 9 US-09-815-242-13768 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

6 1155 32.5 734 9 US-09-815-242-10312 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

7 1134 31.9 709 9 US-09-815-242-10989 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

8 1132 31.8 705 11 US-09-907-907A-43 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

9 1128.5 31.7 704 9 US-09-815-242-10882 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

10 1127 31.7 702 9 US-09-815-242-4975 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

11 1097 30.8 613 9 US-09-815-242-11755 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

12 1097 30.8 701 9 US-09-815-242-12050 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

13 1054 29.6 737 9 US-09-815-242-13699 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

14 104.5 28.5 753 10 US-09-738-246-5665 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

15 999.5 28.1 738 15 US-10-156-761-10060 Sequence 42, Appl Sequence 44, Appl Sequence 543, Ap Sequence 12325, A Sequence 13768, A Sequence 10312, A Sequence 10989, A Sequence 43, Appl Sequence 1082, A Sequence 4975, A Sequence 11755, A Sequence 12050, A Sequence 13699, A Sequence 5665, Ap Sequence 10060, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-09-907-907A-42
Sequence 42, Application US/0907907A
Publication No. US200300996601
GENERAL INFORMATION:
APPLICANT: Fisher, Paul B.
APPLICANT: Lesczyncka, Magdalena
TITLE OF INVENTION: GENES DISPLAYING ENHANCED EXPRESSION DURING CELLULAR SENESCENCE AND TITLE OF INVENTION: TERMINAL CELL DIFFERENTIATION AND USES THEREOF
FILE REFERENCE: A34584-A-PCT-USA (070005.1664)
CURRENT APPLICATION NUMBER: US/09/907,907A
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: US 09-243,277
PRIOR FILING DATE: 1999-02-02
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 42
LENGTH: 705
TYPE: PRT
ORGANISM: Homo sapiens
US-09-907-907A-42

Query Match 100.0%; Score 3557; DB 11; Length 705; Best Local Similarity 100.0%; Pred. No. 3, 4e-302; Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGFFLLPDRRALTQLQVRA.....TAVLLHNTQLDNERLNILP 705
DQ 1 DGFFLLPDRRALTQLQVRA.....TAVLLHNTQLDNERLNILP 705

Sequence 252, Ap Sequence 11416, A Sequence 45, Ap Sequence 11576, A Sequence 1451, Ap Sequence 1115, Ap Sequence 1533, Ap Sequence 455, Ap Sequence 10721, Ap Sequence 6245, Ap Sequence 1998, Ap Sequence 10287, Ap Sequence 4977, Ap Sequence 7234, Ap Sequence 18348, Ap Sequence 12104, Ap Sequence 664, Ap Sequence 998, Ap Sequence 10481, Ap Sequence 106, Ap Sequence 16635, Ap Sequence 1220, Ap Sequence 1168, Ap Sequence 13393, Ap Sequence 3, Ap

QY 181 GEYVNUPTRKEMSSSTLNULVAGAPKSQIUMLESAENLQDCHAIKVKGVKYQDQ 240
 Db 181 GEYVNUPTRKEMSSSTLNULVAGAPKSQIUMLESAENLQDCHAIKVKGVKYQDQ 240
 QY 241 GIQQLVKETGVTKRTPOQLFTPSBPKVTHKLAMERLYAVFTYEDKVSDEAVNKIR 300
 Db 241 GIQQLVKETGVTKRTPOQLFTPSBPKVTHKLAMERLYAVFTYEDKVSDEAVNKIR 300
 QY 301 LDTEBOLKEKPREADPYTIESFNTVAKVERSVIYNEYKRCRDLTSIRNVSCEVDMF 360
 Db 301 LDTEBOLKEKPREADPYTIESFNTVAKVERSVIYNEYKRCRDLTSIRNVSCEVDMF 360
 QY 301 LDTEBOLKEKPREADPYTIESFNTVAKVERSVIYNEYKRCRDLTSIRNVSCEVDMF 360
 Db 361 KTLGSAFQRGQTQVLTCTVTPDSLESGKSDQVITANGIKDNFMLHYEPYATNEI 420
 Db 361 KTLGSAFQRGQTQVLTCTVTPDSLESGKSDQVITANGIKDNFMLHYEPYATNEI 420
 QY 421 GKVTLNRLRGHGALEAKALYVPIPRDPFTIRTSEVLENSGSSMASACGSSLAMD 480
 Db 421 GKVTLNRLRGHGALEAKALYVPIPRDPFTIRTSEVLENSGSSMASACGSSLAMD 480
 QY 481 SGVPISSAVAGVAIGLVKTDPEKEBIDYRLLTDIGEDYDNGDMDFKLAGTNKGITAL 540
 Db 481 SGVPISSAVAGVAIGLVKTDPEKEBIDYRLLTDIGEDYDNGDMDFKLAGTNKGITAL 540
 QY 541 QADIKLPGIPIKIVMEAOQASVAKKEIQLQIMKTTSKPRASRKENGPVETVQPLSKR 600
 Db 541 QADIKLPGIPIKIVMEAOQASVAKKEIQLQIMKTTSKPRASRKENGPVETVQPLSKR 600
 QY 601 AKFVPGGYNLKLQOAGTGTISQDDEFTSFVAPTPSMHEARDFTICICKDQEQQLE 660
 Db 601 AKFVPGGYNLKLQOAGTGTISQDDEFTSFVAPTPSMHEARDFTICICKDQEQQLE 660
 QY 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 Db 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 QY 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 Db 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 RESULT 2
 US-09-907-907A-44
 ; Sequence 44, Application US/09907907A
 ; Publication No. US20030099660A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fisher, Paul B.
 ; APPLICANT: Leszczyniecka, Magdalena
 ; TITLE OF INVENTION: GENES DISPLAYING ENHANCED EXPRESSION DURING CELLULAR SENESCENCE A
 ; FILE REFERENCE: A34594-A-PCT-USA (07/05/2001)
 ; CURRENT APPLICATION NUMBER: US/09/907,907.
 ; CURRENT FILING DATE: 2001-07-16
 ; PRIOR APPLICATION NUMBER: US 09/243,277
 ; PRIOR FILING DATE: 1999-02-02
 ; NUMBER OF SEQ ID NOS: 51
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 44
 ; LENGTH: 705
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ;
 Query Match 100.0%; Score 3557; DB 11; Length 705;
 Best local similarity 100.0%; Pred. No. 3.4e-302;
 Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGPFLPDRDRLTOLQVRLWSSAGSRRAVADLGNRKLKESGKLRPADGSAWQGD 60
 Db 1 DGPFLPDRDRLTOLQVRLWSSAGSRRAVADLGNRKLKESGKLRPADGSAWQGD 60
 QY 61 TAVMTAVSKTKPSQFPLWVYQKAAAGKPTNVRREVTSKELTSRIDS 120
 Db 61 TAVMTAVSKTKPSQFPLWVYQKAAAGKPTNVRREVTSKELTSRIDS 120
 QY 121 TPLPFPAGFYTQVLNCNLAVDGVNEPDVLAINGASVALSLSDIPWNGPVGAVRIGID 180
 ;
 Db 121 IRLPFPAGFYTQVLNCNLAVDGVNEPDVLAINGASVALSLSDIPWNGPVGAVRIGID 180
 QY 181 GEYVNUPTRKEMSSSTLNULVAGAPKSQIUMLESAENLQDCHAIKVKGVKYQDQ 240
 Db 181 GEYVNUPTRKEMSSSTLNULVAGAPKSQIUMLESAENLQDCHAIKVKGVKYQDQ 240
 QY 241 GIQQLVKETGVTKRTPOQLFTPSBPKVTHKLAMERLYAVFTYEDKVSDEAVNKIR 300
 Db 241 GIQQLVKETGVTKRTPOQLFTPSBPKVTHKLAMERLYAVFTYEDKVSDEAVNKIR 300
 QY 301 LDTEBOLKEKPREADPYTIESFNTVAKVERSVIYNEYKRCRDLTSIRNVSCEVDMF 360
 Db 301 LDTEBOLKEKPREADPYTIESFNTVAKVERSVIYNEYKRCRDLTSIRNVSCEVDMF 360
 QY 301 LDTEBOLKEKPREADPYTIESFNTVAKVERSVIYNEYKRCRDLTSIRNVSCEVDMF 360
 Db 361 KTLGSAFQRGQTQVLTCTVTPDSLESGKSDQVITANGIKDNFMLHYEPYATNEI 420
 Db 361 KTLGSAFQRGQTQVLTCTVTPDSLESGKSDQVITANGIKDNFMLHYEPYATNEI 420
 QY 421 GKVTLNRLRGHGALEAKALYVPIPRDPFTIRTSEVLENSGSSMASACGSSLAMD 480
 Db 421 GKVTLNRLRGHGALEAKALYVPIPRDPFTIRTSEVLENSGSSMASACGSSLAMD 480
 QY 481 SGVPISSAVAGVAIGLVKTDPEKEBIDYRLLTDIGEDYDNGDMDFKLAGTNKGITAL 540
 Db 481 SGVPISSAVAGVAIGLVKTDPEKEBIDYRLLTDIGEDYDNGDMDFKLAGTNKGITAL 540
 QY 541 QADIKLPGIPIKIVMEAOQASVAKKEIQLQIMKTTSKPRASRKENGPVETVQPLSKR 600
 Db 541 QADIKLPGIPIKIVMEAOQASVAKKEIQLQIMKTTSKPRASRKENGPVETVQPLSKR 600
 QY 601 AKFVPGGYNLKLQOAGTGTISQDDEFTSFVAPTPSMHEARDFTICICKDQEQQLE 660
 Db 601 AKFVPGGYNLKLQOAGTGTISQDDEFTSFVAPTPSMHEARDFTICICKDQEQQLE 660
 QY 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 Db 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 QY 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 Db 661 FGAVYATTITEIRDGVWVKLYPNMTAVLHNTOLDNERNLILP 705
 RESULT 3
 US-09-815-242-5443
 ; Sequence 543, Application US/09815242
 ; Patent No. US2002001569A1
 ; GENERAL INFORMATION:
 ; Sequence 543, Application US/09815242
 ; APPLICANT: Haebeck, Robert
 ; APPLICANT: Olsen, Kari L.
 ; APPLICANT: Zybird, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Travich, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: BLTRA_01A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ;
 Query Match 100.0%; Score 3557; DB 11; Length 705;
 Best local similarity 100.0%; Pred. No. 3.4e-302;
 Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGPFLPDRDRLTOLQVRLWSSAGSRRAVADLGNRKLKESGKLRPADGSAWQGD 60
 Db 1 DGPFLPDRDRLTOLQVRLWSSAGSRRAVADLGNRKLKESGKLRPADGSAWQGD 60
 QY 61 TAVMTAVSKTKPSQFPLWVYQKAAAGKPTNVRREVTSKELTSRIDS 120
 Db 61 TAVMTAVSKTKPSQFPLWVYQKAAAGKPTNVRREVTSKELTSRIDS 120
 QY 121 TPLPFPAGFYTQVLNCNLAVDGVNEPDVLAINGASVALSLSDIPWNGPVGAVRIGID 180
 ;

; SEQ ID NO 5443
; LENGTH: 696
; TYPE: PRT
; ORGANISM: *Staphylococcus aureus*
; US-09-815-242-5443

TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: E11TRA.01A
CURRENT APPLICATION: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078

PRIORITY FILING DATE: 2000-01-21

PRIORITY APPLICATION NUMBER: 60/205,848

PRIORITY FILING DATE: 2000-05-23

PRIORITY APPLICATION NUMBER: 60/207,727

PRIORITY FILING DATE: 2000-05-26

PRIORITY APPLICATION NUMBER: 60/242,578

PRIORITY FILING DATE: 2000-10-23

PRIORITY APPLICATION NUMBER: 60/253,625

PRIORITY FILING DATE: 2000-11-27

PRIORITY APPLICATION NUMBER: 60/257,931

PRIORITY FILING DATE: 2000-12-22

PRIORITY APPLICATION NUMBER: 60/269,308

PRIORITY FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 1410

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 1225

LENGTH: 698

TYPE: PRT

ORGANISM: *Staphylococcus aureus*

Query Match 33.5%; Score 1191.5; DB 9; Length 696;
Best Local Similarity 39.2%; Pred. No. 4.1e-95; Matches 274; Conservative 117; Mismatches 249; Indels 59; Gaps 12;
; Db 24 SAGSRAVAVDGNRKLEISSGKLRPADGSAVQSGDTAVMTAVSKTKPSQFQMLVV 83
; Db 2 SOBKVKVFTTEWAGRSLTETGQLAKOANGAVAVLRYGDTWVLSTATASKRPRDGDFPPLTV 61
; Db 84 DYROKAAAGRIPTNVLREVGTSKDEBTSRIDSIRPFPAGYFYDQVLCNLIAVD 143
; Db 62 NYEKVAAAGKIPGPKRGRPGDATIARLIRPIRLFLPKYKHDQIMMNLSD 121
; Db 144 GVNEDPVLAINGASVSLSDIPWNGPGAVARGIGIDGEBYVNPTRKEMSSSTLNUVAG 203
; Db 122 PDCSPOMAAMIGSSMALSVDIFQGPIAGVNGVYDGYKTYINPTVEKEVSRLLEVAG 181
; Db 204 ARPKSQVMLESAENILQDQFCHAIVKG-----VKYTOQIQGQOLVKEVTKRTP 256
; Db 182 -HKDAVNVMVAGASBITEQMLEAIFFGHEBIRQLVDFOQIVDHIOPVKOE----- 232
; Db 257 QKLFTPS-----PRVKYTHKLAMERLYAVFTYEDKVSRSRDAVNRKIRLDTEQQLKE 309
; Db 233 --FIPADERBALVERIKSLSLTETGQLAKOANGAVAVLRYGDTWVLSTATASKRPRDGDFPPLTV 301
; Db 310 KFPEADP-----YE11ESFNWVAKEVFRSIVLNERYKCDGRDLTSRANVSCYDMPKTL 363
; Db 282 FIDEDEPENELLIKEVYAILNELVKEVRRLIADEKIRPDKRKPDEIRPLDSEVGILPRT 341
; Db 364 HGSALFORQGQVQLCITWFDSSLESGIKSDQVITAINGIKDNFMHLHYEPYATNEIGKV 423
; Db 342 HGSGLFTRGQTOALSVTIGAL-----GDXQUDLGLGPEERKRPMHYNFPNFSVGTPV 397
; Db 424 TGLNRBLGIGALAEKALYVIP-----RDPPTTRVTSVLESLNGSSMASACGSLALMDS 481
; Db 458 GVPKAPVAGIAMGLVRED-----SYTILTDIOMEDALGDMDFKVAQPTKEGITAQ 510
; Db 398 RAPGRREIGHGALAGRAKVIIDETGVKUDIEQDGTFIFGAVDQMINRAREIEITR-----EAEV 625
; Db 511 MDIKIDGLTRTEIEALFQARRGRLEIIMNHMQTIDOPRTLSAYAPKQVUMTIDKIR 570
; Db 542 ADIKLPGIPKIVMELQOSVAKKELIQQIMKTTISKPRASKRKENGENGVETVQPLSKRA 601
; Db 510 KFPEADP-----YE11ESFNWVAKEVFRSIVLNERYKCDGRDLTSRANVSCYDMPKTL 633
; Db 571 DVIGPGKKKINELIDETGVKUDIEQDGTFIFGAVDQMINRAREIEITR-----EAEV 625
; Db 662 GAVYATATITEUTGVWVVKLYNNMTAVLHQNOLDNRL 700
; Db 626 GOTYQATVKRICKYGAFTVGLFPKDA-LIHSIQSKNRI 663

Query Match 33.5%; Score 1191.5; DB 9; Length 698;
Best Local Similarity 39.2%; Pred. No. 4.1e-95; Matches 274; Conservative 117; Mismatches 249; Indels 59; Gaps 12;
; Db 24 SAGSRAVAVDGNRKLEISSGKLRPADGSAVQSGDTAVMTAVSKTKPSQFQMLVV 83
; Db 2 SOBKVKVFTTEWAGRSLTETGQLAKOANGAVAVLRYGDTWVLSTATASKRPRDGDFPPLTV 61
; Db 84 DYROKAAAGRIPTNVLREVGTSKDEBTSRIDSIRPFPAGYFYDQVLCNLIAVD 143
; Db 62 NYEKVAAAGKIPGPKRGRPGDATIARLIRPIRLFLPKYKHDQIMMNLSD 121
; Db 144 GVNEDPVLAINGASVSLSDIPWNGPGAVARGIGIDGEBYVNPTRKEMSSSTLNUVAG 203
; Db 122 PDCSPOMAAMIGSSMALSVDIFQGPIAGVNGVYDGYKTYINPTVEKEVSRLLEVAG 181
; Db 204 ARPKSQVMLESAENILQDQFCHAIVKG-----VKYTOQIQGQOLVKEVTKRTP 256
; Db 182 -HKDAVNVMVAGASBITEQMLEAIFFGHEBIRQLVDFOQIVDHIOPVKOE----- 232
; Db 257 QKLFTPS-----PRVKYTHKLAMERLYAVFTYEDKVSRSRDAVNRKIRLDTEQQLKE 309
; Db 233 --FIPADERBALVERIKSLSLTETGQLAKOANGAVAVLRYGDTWVLSTATASKRPRDGDFPPLTV 301
; Db 310 KFPEADP-----YE11ESFNWVAKEVFRSIVLNERYKCDGRDLTSRANVSCYDMPKTL 363
; Db 282 FIDEDEPENELLIKEVYAILNELVKEVRRLIADEKIRPDKRKPDEIRPLDSEVGILPRT 341
; Db 364 HGSALFORQGQVQLCITWFDSSLESGIKSDQVITAINGIKDNFMHLHYEPYATNEIGKV 423
; Db 342 HGSGLFTRGQTOALSVTIGAL-----GDXQUDLGLGPEERKRPMHYNFPNFSVGTPV 397
; Db 424 TGLNRBLGIGALAEKALYVIP-----RDPPTTRVTSVLESLNGSSMASACGSLALMDS 481
; Db 398 RAPGRREIGHGALAGRAKVIIDETGVKUDIEQDGTFIFGAVDQMINRAREIEITR-----EAEV 625
; Db 482 GVPITSAVAGVAIGLVTKTPBKGEBIYDRLTDIGTIEDYNGDMDFKLAGTNKGITALQ 541
; Db 458 GVPKAPVAGIAMGLVRED-----SYTILTDIOMEDALGDMDFKVAQPTKEGITAQ 510
; Db 542 ADIKLPGIPKIVMELQOSVAKKELIQQIMKTTISKPRASKRKENGENGVETVQPLSKRA 601
; Db 511 MDIKIDGLTRTEIEALFQARRGRLEIIMNHMQTIDOPRTLSAYAPKQVUMTIDKIR 570
; Db 542 KFVGQGGYNKLQAEQETGVTSQDVDEETFSVAPPSSVMHEARDRTECKDDQEQQLEF 661
; Db 602 KFVGQGGYNKLQAEQETGVTSQDVDEETFSVAPPSSVMHEARDRTECKDDQEQQLEF 661
; Db 571 DVIGPGKKKINELIDETGVKUDIEQDGTFIFGAVDQMINRAREIEITR-----EAEV 625

RESULT 4
US-09-815-242-12325
; Sequence 12325, Application US/09815242
; Patent No. US20020206156A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

QY 662 GAVYATTETRDTGVMVKLYPNMPTAVLINTQLDNERL 700
 ; Sequence 13768, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hasebeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; Title of Invention: Prokaryotes
 ; FILE REFERENCE: ELITRA.01A
 ; CURRENT APPLICATION NUMBER: US/09/815, 242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191, 078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206, 848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207, 727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242, 578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253, 625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257, 931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269, 308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 13768
 ; LENGTH: 721
 ; TYPE: PRT
 ; ORGANISM: Salmonella typhi

QY NAME/KEY: VARIANT
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 ; US-09-815-242-13768

Query Match 32.6%; Score 1158; DB 9; Length 721;
 Best Local Similarity 39.9%; Pred. No. 3.7e-92;
 Matches 270; Conservative 117; Mismatches 250; Indels 40; Gaps 11;

QY 35 GNRKURISGKLLARFADGSWVQSGDTAVMTAVSKTKPSQ-FMLPVYDYRKAAG 93
 ; Sequence 10312, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hasebeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; Title of Invention: Prokaryotes
 ; FILE REFERENCE: ELITRA.01A
 ; CURRENT APPLICATION NUMBER: US/09/815, 242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191, 078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206, 848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207, 727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242, 578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253, 625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257, 931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269, 308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 10312
 ; LENGTH: 734
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; US-09-815-242-10312

Query Match 32.5%; Score 1155; DB 9; Length 734;
 Best Local Similarity 39.9%; Pred. No. 6.9e-92;
 Matches 269; Conservative 118; Mismatches 252; Indels 36; Gaps 11;

QY 214 ASAENILQODFCHAKVKGVKQTQIGIQLQVKETGVTKRTPKLFTSPETK---- 268
 ; Sequence 201, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hasebeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; Title of Invention: Prokaryotes
 ; FILE REFERENCE: ELITRA.01A
 ; CURRENT APPLICATION NUMBER: US/09/815, 242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191, 078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206, 848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207, 727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242, 578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253, 625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257, 931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269, 308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 10312
 ; LENGTH: 734
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; US-09-815-242-10312

Query Match 32.5%; Score 1155; DB 9; Length 734;
 Best Local Similarity 39.9%; Pred. No. 6.9e-92;
 Matches 269; Conservative 118; Mismatches 252; Indels 36; Gaps 11;

PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 10899
 LENGTH: 709
 TYPE: PRT
 ORGANISM: *Haemophilus influenzae*
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1): (709)
 OTHER INFORMATION: xaa = Any Amino Acid
 US-09-815-242-10989
 Query Match 31.9%; Score 1134; DB 9; Length 709;
 Best Local Similarity 36.8%; Pred. No. 4.5e-90;
 Matches 249; Conservative 137; Mismatches 251; Indels 40; Gaps 12;
 Db 35 GNRKLEISSGKLLARPADGSAVQSGDTAVMTAVSKTKPSPSQ-FMPLVVDYRQKAAG 93
 Db 35 GQHTVLTETGMMARQATAVWMSMDTAVFVTWVQGQKAKCQDFPLTVWQERTVAG 94
 Qy 94 RIPTNLYREVGSDKETLTSRITDRSTRPLPAGYFYDVTQVLCNLAVD3YNEDPLAI 153
 Db 95 RIGSFRRERGRPSGEETLARLIDRPIRPLPPEGFVNEVQIATVSVNPOVNPDIVAM 154
 Qy 154 NGASVALSLSDIPWNPYGAVARIGIDGEYVNPTRKMSSTNLVAGPKSQIVMLE 213
 Db 155 IGASAALSLSGIFNGPPIGAARYVINDQVWAMPDKRKLIVWAGT-EAAYLNE 213
 Db 214 ASAENILQODFCHAIKVQVKTQOIQIQLQVKETGVTKTPQKLTSP--EIVKVT 270
 Db 214 SERQOLLSBDQMLGAWVFGHEQQPVQVGNINELVKEAG---KPRWQKOPPNEALNARV 269
 Qy 271 HKLAMERL--YAVFTDVEHDKVSRSDEAVNKRLDTRBQLKKEKPFPEADPYEIESNVVA 327
 Db 270 AIAEAEALSDAVRI---TDKQERYAOVWIKSETATLADEDTDENEIGEILHAE 324
 Qy 328 KEVFRSTVLINEYKRCRDGLTSRNLNVACEDVMDFTKLGSLALQORGQVQICVTFDLES 387
 Db 325 KTVRSVRLGGERPRIDEKDKMRGLDRTGVLPRTRGSAITGTRGETOALVATLGN-- 381
 Qy 388 GIKSQVITAINGIKDKNFMLYEPPYATNEIGKVTLNRLGHGALAKALYVIP- 446
 Db 382 -ARDAQVILDELMGERTDTPLFHNFPPVSYGETGMGSPKREIGHLRGRVLAUMPD 440
 Qy 447 -RDPFTTIRVTEVLENSGSSMASACCGSISLAMDSCVGPISAVAGVAIGVLTQKDPKG 505
 Db 441 MDKEPYTWRVVBSEITSGSSMASVCGASLALMDAGVPIKAVAGIATAMGLVKEGD--- 496
 Qy 506 EIDYRLITDIDGIEDYNGMDPKIAGTNGKITALQADIKLPGIPKIKVIMAIQOSVAK 565
 Db 497 --NVVVLSDILGDEDHGDMDPKVASSRDRGJSALQMDIKRIGEITKIMQVANQKAR 553
 Qy 566 KELQIMKTTSKPRAKRKGKENGPPVETVQVPLSKRAKFVGGYNNKKLQAEVTQISQV 625
 Db 554 LHLIGWMEQANAPRGDSEFAPRINTKINDKIKOVIGKGSVIRALTETGTTIE 613
 Qy 626 DBTFSVFAPTPSVMHEARDFTIECKDQDQOQLERGAVYATITTEIRDGVNWKLYPNM 685
 Db 614 DDTGKVKIAATGEKAKHAIARRIEE-----AEIEVERGVYIGKVTRIVDFGAFVAIGGGK 668
 Qy 686 TAVLHNTQLDIWERL 700
 Db 669 EG-LVHISQIADKRV 682
 RESULT 7
 US-09-815-242-10989
 ; Sequence 10989, Application US/09015242
 ; Patent No. US2002006156A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haelbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA_01A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIORITY FILING NUMBER: 60/191,078
 ; PRIORITY FILING DATE: 2000-03-21
 ; PRIORITY APPLICATION NUMBER: 60/206,848
 ; PRIORITY FILING DATE: 2000-05-23
 ; PRIORITY APPLICATION NUMBER: 60/207,727

Qy 391 SD-QVITAINGIKDKNFMLYEPPYATNBIGKVTLNRLGHGALAKALYVIP--R 447
 Db 359 RDAQITDELTGERQDHFLFNFPPVSYGETGMGSPKREIGHLRGRVLAUMPLA 418
 Qy 448 DFPFTTIRVTEVLENSGSSMASACCGSISLAMDSCVGPISAVAGVAIGVLTQKDPKG 507
 Db 449 EFPVYVWRVSEITSGSSMASVCGASLALMDAGVPIKAVAGIATAMGLVK-----EE 471
 Qy 508 EIDYRLITDIDGIEDYNGMDPKIAGTNGKITALQADIKLPGIPKIKVIMAIQOSVAKKE 567
 Db 472 EKFLVLSDILGDEDHGDMDPKVAGTREGTALQMDIKRIGEITKIMQVANQKARH 531
 Qy 568 ILQIMKTTSKPRAKRKGKENGPPVETVQVPLSKRAKFVGGYNNKKLQAEVTQISQV 627
 Db 532 LHLIGWMEQANAPRGDSEFAPRINTKINDKIKOVIGKGSVIRALTETGTTIE 591
 Qy 628 BTFSVFA---PTPSVMHEARDFTIECKDQDQOQLERGAVYATITTEIRDGVNWKLYP 683
 Db 592 GTVKIAAVDSNAKVMGRIBETVAEV-----EAGVYIGKVTRIVDFGAFVAIG 642
 Qy 684 NMATVLLHNTQLDIWERL 700
 Db 643 NKEG-LVHISQIADKRV 658

Db 340 RYHGSGIPTRGPTQASVCTL---APLGEMOIIDGL-GVQDSKRFIHHMVPQPSVGST 394
 Qy 421 GKVGTGLNRELGALGAKALYKPVIP--RDPFTTIRVSEVLESNGSSMSACGSSLAL 478
 Db 395 GRAGSPGEREIGHGALGERALQIQIPSEEDPFPTIRLVAEVLESNGSSSQASICAGTL 454
 Qy 479 MSGVPISSAVAGVAGLGLVTKDPERKEIYDRLDILGEDYNGDMDFKLAGINKGIT 538
 Db 455 MDAGVPIKAPVAGIAGLVS-----DGENYTILTDIQLEDHLGMDMDFKVAGTKDGT 507
 Qy 539 ALQADIKLPGPKIVMIAQASVAKKEIQLQMNKTKSKRASRKGNGPVPETVOPLS 598
 Db 508 ALQMDIKIQLGTTQIETALDQAKKARMEHIELLTIAAREELSQYAPKLEMQIKPA 567
 Qy 599 KRAKFPVPGGTYMKKQOETSYTISQVDEETPSVAPTPSVMHEARDFTICKDQEQQ 658
 Db 568 KKDVKIGKGETINSTIDBTGVKIDIDQGNSIASSDAEMIKKAIKIEBLTK----E 622
 Qy 659 LFGAVVTTATTEIRDGVMVKLYPNTAVLHNTQDNEULN 701
 Db 623 VEVGQVYLAKVRIEKFGAFVNLIKGDG-LHISOLANERVN 664

RESULT 10
 US-09-815-242-4975
 ; Sequence 4975, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; LENGTH: 702
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-815-242-4975

Query Match 31.7%; Score 1127; DB 9; Length 702;
 Best Local Similarity 37.4%; Pred. No. 1.8e-89;
 Matches 254; Conservative 144; Mismatches 242; Indels 40; Gaps 13;

Db 70 IPGGFKREGPSEARLTARLIDRPRPSEGFRNREVOINTIVMSVEDCPTPMAAMP 129
 Qy 155 GASVIALSLSDPWNGPVGAVRIGIDGEGYVNPTEKEMMSSTNLWAGPKSOTMILEA 214
 Db 130 GSSLALAIISDIPFDGPIAGVGVGRINGEVYLNPTVEQAOEDTBLVAGT-KEA1NMVES 188
 Qy 215 SAENLIOQDQCHAIVKVGKTOQIQLQOLVKETGVTQPKQULFSPBEIVKVTKLA 274
 Db 189 GAKEVSEEDMGLGALLFGFDKELVAFQELVAAYGKPKDVL----QVADLKEI 243
 Qy 275 MERLY----AVFTDVEHDKVSREBAVKRLDTEBQKKEKPE-ADPYETIESFNVVA- 327
 Db 244 RDAYTMKTAVMTB--EKLAREVEIDKVQDVTKEVYAKFAEHEBEAQOLKEVVKQIA 300
 Qy 328 ---KEVFRSTVLANEKRCDCRDLSLRNVSCEVDMPKTLGSALFQRGQVLCWTFDS 384
 Db 301 DLKDVKVRELLTIDKRPDKRLDDETRHLSSEVSILPVRHGSGLFTRGPTQALSVTL-- 358
 Qy 385 LESGKSDQVTAIMKID-KNFMLYHEFPYATBEGKVTGUNRRELGHGALAKALY 443
 Db 359 --APLGEMOIIDGL-GVQDSKRFIHHMVPQPSVGSTGRASPSREIGHGALRAQ 415
 Qy 444 VIP--RDPFTTIRVSEVLESNGSSMSACGSSLALMMSGVPISSAVAGLGLVTKD 501
 Db 416 LIPSEDFPPTIRLVAEVLESNGSSSQASICAGTLALMDAGVPIKAPVAGIAGLVS-- 472
 Qy 502 PEKGTEDEYMLTDIPTGIEDYNGDMDFKLAGTNGKITAQDILKLPGIPKIVMIAQQA 561
 Db 473 ---DGENYTILTDIQLEDHLGMDMDFKVAGTKDGTQDIALQMDIKIQLGTTQIET 528
 Qy 562 SVAKELIQLMNKTKSKRASRKGNGPVPETVQVLSKRAKFPVPGGTYMKKQOETGV 621
 Db 529 KKARMEHIELLTIAAREELSQYAPKLEMQIKPAKTDVKIGKGETINSTIDBTGV 588
 Qy 622 ISQVDEETPSVAPTPSVMHEARDFTICKDQDQEQQLBGAVVTTATTEIRDGVMVKL 681
 Db 589 IDIDQGNSIASSDAEMIKKAIKIEBLTK----EVENVQVYLAKVRIEKFGAFVN 643
 Qy 682 YPNMTRAVLHNTQDNEULN 701
 Db 644 IKGDKD-LHISOLANERVN 662

RESULT 11
 US-09-815-242-11755
 ; Sequence 11755, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.01A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; LENGTH: 702
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-815-242-4975

Query Match 31.7%; Score 1127; DB 9; Length 702;
 Best Local Similarity 37.4%; Pred. No. 1.8e-89;
 Matches 254; Conservative 144; Mismatches 242; Indels 40; Gaps 13;

Qy 35 GRKRUKEISSGKLARFADGSAVQSGDTAVMTAVSKTPSPSQFMPFLVUDYRQKALAGR 94
 Db 10 SCRPLEVEIGQALQOANGAVIUYRGTWILVLAASKEAKDVEFFLTUNVEEKOMYAVGK 69
 Qy 95 IPTNYLREVEFTSDKELTSRITDRSTRPLTPAGYFDTQYLCNLIAVDGVNEPDYLAIN 154

QY 683 PNTAVLHLNTQLDNERLN 701
 US-09-815-242-13699
 ; Sequence 13699, Application US/09815242
 ; Patent No. US20020061509A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identifcation of Essential Genes in
 ; FILE REFERENCE: ELITRA_011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: PatnSeq for Windows Version 4.0
 ; SEQ ID NO: 13699
 ; LENGTH: 737
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 ; US-09-815-242-13699

Query Match 29.6%; Score 1054; DB 9; Length 737;
 Best Local Similarity 34.6%; Pred. No. 4.9e-83;
 Matches 234; Conservative 148; Mismatches 257; Indels 38; Gaps 9;

QY 37 RKLRLISGGKLRKFADGSAVVGSGDTAVMTAVSKTKPSKPSQMPVLDYDROKAAAGRIP 96
 Db 13 RBLIVETGQVAKQANGSVVVRIGESTVTLTAIVMSKRMATGDPFPQLQVNEBVKYVAGKFP 72
 QY 97 TNYLRLRVEGTSKELTSRIDSIRPLPAGYFYDQVQNLALTDGVNEDVPAVLAINGA 156
 Db 73 GCFMKRERGRPSDTATLRLDRPIRMPAFGRNEVQVINTVLSDENASAPMAAMFGS 132
 QY 157 SVALSUDPMPGPVGAWRIGLIDGEYVNPTRKEMSSSTNLVAGPKSQIVMELASA 216
 Db 133 SVALSUDPMPGPVGAWRIGLIDGEYVNPTRKEMSSSTNLVAGPKSQIVMELASA 191
 QY 217 ENLQODPCHATKVGKVTQOIQIGQOLVKEGTVKTRTPQKLFPSPEITKTYTHKLA 276
 Db 192 KELSEBIMLEALKGHEBAKELIAFQBEIVAAVGKEK-----AEVELLHVDAEIQAE 243
 QY 277 RLYAVFVDE---RDKVSRDPAVAKNRLDPEQQLKEKPFADPY---EILSFNVV 326
 Db 244 IIAAVNSDLOKAVQVEEKLAETAOQKVQDVTAVKEVYANHEEDRIMDVABLEQ 303
 QY 327 AKEVFRSIVLNUYKRCGDLTSLRNVSCEUDMFKTLHGSALFORGOTQVCLVCTVFDSE 386
 Db 304 EHAEVRLITEDKVRPDKRKYDSEIRPLDAVVDPLPRVHGSGUFTRGOTQALSVLTU--- 359

RESULT 13
 US-09-815-242-13699
 ; Sequence 13699, Application US/09815242
 ; Patent No. US20020061509A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identifcation of Essential Genes in
 ; FILE REFERENCE: ELITRA_011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: PatnSeq for Windows Version 4.0
 ; SEQ ID NO: 13699
 ; LENGTH: 737
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 ; US-09-815-242-13699

RESULT 14
 US-09-738-626-5665
 ; Sequence 5665, Application US/09738626
 ; Publication No. US20020197305A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738, 626
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO: 5665
 ; LENGTH: 753
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-5665

Query Match 28.5%; Score 1014.5; DB 10; Length 753;
 Best Local Similarity 35.8%; Pred. No. 1.4e-79;
 Matches 255; Conservative 125; Mismatches 263; Indels 69; Gaps 19;

QY 33 DLGNRKEISGGKLRKFADGSAVQG-SGPTAVMVTAVSKTKPSPS-QEMPLVUDYRQKA 90
 Db 25 DFGTRTRFEGQLARQADGAVTVDLDTMLATTTASHNQPREGFDFFPLTVDTERM 84
 QY 91 AAGRPTNLYRVEGTSKELTSRIDSIRPLPAGYFYDQVQNLALTDGVNEDVPAVLA 150
 Db 85 AAGRPTNLYRVEGTSKELTSRIDSIRPLPAGYFYDQVQNLALTDGVNEDVPAVLA 144
 QY 151 LAINGASVALSLSDIPMPGPVGAWRIGI-----DGEYVNPTRKEMMSSTNLVAGA 204

QY 93 GRIFTINLIRENGTSDEKILTSRIDSIRELPFACYFDIQUICLNALANGNEPDVIA 152
 205 PKSO-----IVMLEASA-ENILOQ-----DFCHAIVGKVYKTOQI 238
 QY 78 GKPQGTSFRERKRSBDSLTCRCLDRPLRSFFKKGRLNBTQVATIMALPDPHLYDVA 137
 Db 205 LVERKRGKNTPSDVAVMVTEAGASENVVRVKDGAAPTEKIVSDGLEAKPKFDILCRA 264
 QY 153 INCASVALSLSDLPWNCPVGAVRIGIDGEYVNPTRKEMSSSTLNUVAGPKSQ--- 208
 239 1Q3IQOLVKEGTGVTRTPOGLFT-SPEVTKYTHKLAMERLYVFTDYEHDYKVSRSDEAVN 297
 QY 138 INVASACTQDAGLPFSCOPIGGVVALINGQWAPTFTEEDAVFDMVAGBAGEDVA 197
 Db 265 OEGJLQRV--GNAAMEKFP--LFPPYTDDEVSAVERKVSKKLASSLT--LKAKOBRDDATN 318
 QY 209 IVMLEASERNTIQQDFCHAIVGKVYQOIQGIOLVKE--TGYTKRTS--OKLFTPS 263
 Db 298 KIRLDEEQLEKEF----PEADPYETIESFNVVAKEVFRSTVILNSYKRCGDRDLSL 350
 QY 198 IMVVAEATE----KTIOLVAGGAEAPTEEVVAGLDAAKPDKVILCAQ 243
 Db 319 AYMEBIEASILPKPEASVSSAAEASK-EFRAGYNAVMKAIVRMLTDFHFRDGRGVDI 377
 QY 411 EPPYATNEIGKVGTGLNRELGHALABCALYVPLP-RDPFTIRVSEVLESNGSSM 468
 Db 351 RNVSCEDVMPKTLIGSALFORQOTVLCWTFPSLESGTSKSDQVITAINGIKONFMFLHY 410
 QY 378 RDLAEVEELPRAHGSLSLFRGERETQIQLGVTTLDML---KMEQOIDSJAPGDAKRYMHY 433
 Db 431 RPPYATNEIGKVGTGLNRELGHALABCALYVPLP-RDPFTIRVSEVLESNGSSM 468
 QY 434 NPPYSTGTGRVGSKRRBIGHGALAAERAVLPVIPSREBFPYAIRVSEALGSNSGSTM 493
 Db 469 ASRGGGSALMDSGVPISSAVAGVAGLIVGKGE--DYLITDINGDIEDYNGD 525
 QY 548 MDFKVAQTADPITALQDTCUDGIPSKVLSDALEQARDARLTINTMADVINGPDEMSK- 606
 Db 494 GSVCASTLSLYNAGVPLKAPVAGJAMGLVS----GEGDKBEVATDILGADFGD 547
 QY 526 MDPKLAGTNGKIGTALAOADTKLPGSPKIKIIVMEIAQOASVAKKEILOQMKNTISKPRASRKE 585
 Db 586 NGPVEVETQVPLSKRAKEFWPGGYNLKKQQAETGVTISQVDEETFFSVFAPTPSVMHEARD 645
 QY 607 FAPRITVVKIPVAKGELIGPKGKINNATEETGANTSIEDDGTIVFSAADGSAEALIE 666
 Db 646 FITEIKCKDDEQOLFEGAVYATATTEIRDGTGVMWVKLYPNTAVLHNTOLDN 697
 QY 667 KINALA---NPQLPKVGERPLGLTVKTMGAFVSLIPLGDRG-LWHISKIGN 714
 Db
 RESULT 15
 US-10-156-761-10060
 ; Sequence 10060, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAKOSHI
 ; APPLICANT: SAKAKI, YOSHIOUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156, 761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 10060
 ; LENGTH: 738
 ; TVFB: PRT
 ; ORGANISM: Streptomyces avermitilis
 ; US-10-156-761-10060
 Query Match 28.1%; Score 999.5; DB 15; Length 738;
 Best Local Similarity 34.5%; Pred. No. 2.9e-78;
 Matches 240; Conservative 130; Mismatches 262; Indexes 63; Gaps 16;
 QY 35 GNRKLEISGKLLARFDGSVWQ-SGDPAVMVTAWSKTKPSS-QFMPILWVYRQKAAA 92
 QY 18 GTPTRFEFGRLLKQAAQSAVAYDFTWPLSATSKPKPENLDFELTIVVETRMAA 77
 Db

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

DM protein - protein search, using sw model

run on: January 8, 2004, 10:33:56 ; Search time 27 Seconds

(without alignments)

2511.071 Million cell updates/sec

Title: US-09-907-907A-42

Sequence: 1 DGPFLPDRDRLAQVRA. TAVLHNTQLDNERLNLIP 705

Scoring table: BLOSUM62

Gapop: 10.0 , Gapext: 0.5

searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

PIR 76:*
1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

SUMMARIES

result No.	Score	Query Match Length	DB ID	Description
1	1365.5	38.4	991	T48631
2	1244.5	33.9	703	B97123
3	1190.5	33.5	698	B89910
4	1166.5	32.6	713	AE5861
5	1166.5	32.8	713	F97368
6	1161	32.6	711	2 AD0501
7	1158	32.6	734	2 E91134
8	1158	32.6	734	2 H85979
9	1158	32.6	734	2 G75320
10	1156	32.5	810	2 S74509
11	1153	32.4	718	2
12	1152	32.4	734	2 H65106
13	1150	32.3	714	2 AC3497
14	1146	32.2	705	2 AC0424
15	1145	32.2	704	2 G83950
16	1140.5	32.1	745	2 B71654
17	1140.5	32.1	775	2 B70320
18	1138.5	32.0	718	2 AD2355
19	1137	32.0	719	2 F82831
20	1134	31.9	709	2 B64056
21	1132	31.8	705	2 S70691
22	1130	31.8	723	2 AG1603
23	1127	31.7	723	2 AC1241
24	1121	31.5	709	2 E82298
25	1115	31.3	707	2 E84973
26	1114.5	31.3	694	2 D86615
27	1114.5	31.3	694	2 G72009
28	1112	31.3	707	2 C81601
29	1111	31.2	706	2 H81943

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT	1
T48631	polynucleotide phosphorylase - <i>Arabidopsis thaliana</i>
N	Alternate name: protein T15N1.70
C	Species: <i>Arabidopsis thaliana</i> (mouse-ear cress)
C	Date: 20-Apr-2000 #sequence_revision: 20-Apr-2000 #text_change: 20-Apr-2000
C	Accession: T48631
R	Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; submitted to the Protein Sequence Database, April 2000
A	Reference number: 224493
A	Accession: T48631
A	Status: preliminary
A	Molecule type: DNA
A	Residues: 1-991 <BEV>
A	Cross-references: EMBL:AL163792
A	Experimental source: cultivar Columbia; BAC clone T15N1
A	Genetics:
A	Map position: 5
A	Introns: 107/3; 159/3; 216/3; 258/3; 312/3; 352/3; 398/3; 453/1; 495/1; 538/3; 612/1
A	Note: T15N1.70
Query	MPLVDPYKQAAAGRIPINYLREVGTSKDELTSRIDSRTSPRLPAGYFVYQTLCN
Match	MPLVDPYKQAAAGRIPINYLREVGTSKDELTSRIDSRTSPRLPAGYFVYQTLCN
Best	Local Similarity 41.7%; Score 1365.5; DB 2; Length 991;
Matches	287; Conservative 138; Mismatches 231; Indels 33; Gaps 11;
QY	24 SAGSRAV----AVDLGNRKLLEISSGKLARFAGSASVQSGDAAWVMTAVSKTKPSQF
Db	44 SAGSRAV----AVDLGNRKLLEISSGKLARFAGSASVQSGDAAWVMTAVSKTKPSQF
QY	79 MPLVDPYKQAAAGRIPINYLREVGTSKDELTSRIDSRTSPRLPAGYFVYQTLCN
Db	104 MPLVDPYKQAAAGRIPINYLREVGTSKDELTSRIDSRTSPRLPAGYFVYQTLCN
QY	139 LLAVDGVMNPBDVLAINGASVALSLSIDPANGPGAVARGIDSEYYVNPNTPKENSSTLN
Db	164 VLSSDGKQPDIDIANASSAALMSDVMWGGPQVIRIGRIGQFVVNPMTDLSLSSDLN
QY	199 LIVAGAPKSQIVMLAEASENILQDGFCHAIVK---GKVKTOIOGQOLVQVKGTVKR
Db	224 LIVA-CTDRKTMWIDVQREISERKDLAAALRPAEAVKYLDDPI---RLAERAKGQK
QY	255 TPKQLKFTSPPEIVKTHKLAMERLYAVFTDYEHDKVSDAEVVKIRLTDPEEQLEKEKFPPEA
Db	279 -EYKLSMLSDKTLERKVADLAATHTIESVSFIDPSYKFERGEALDNGKVKVFEEGDQE
QY	315 DPPELIESFVNVAKEVFESIIVLNBYKRCGDRITSLRAVSCEDVDMFKTLHGSALFORQ
Db	338 SLSLPLKAVDVTKVKVQVSRMSDGPFVDRGHDVPRVYCESHYLPALHGSALFSGT
QY	375 QVCTVTFDSLESGIKSPDQVITATGKDNFLHYERPPYANNEIGKVTGLURRELHG
Db	398 QVCTVTFDSLESGIKSPDQVITATGKDNFLHYERPPYANNEIGKVTGLURRELHG

QY	435 ALAKKALYPVPIRDL--FPEFTIRVTEBVLESNGSSMASACGGSLALMDSGVPISSAVAGV	492	Db	359 AQVTDGLQYEVESGRYMHYNFPYSTGSKVPLRGPNRRBIGHALAKCALVLPISHEF	418
Db	454 TAKGKALLAVLUPPERBAPPTRINSEVMGSDGTSMASVGGSMALMDAGIPPLRAHVAGV	513	QY	450 PFTIRVTEBVLESNGSSMASACGGSLALMDSGVPISSAVAGVAGLVTKTPPEKGED	509
QY	493 AIGLVTKTPPKGSILEDYUUTDIDGIEDYNGDMDPKINGTKNTGKITALQADIKLPGNIK	552	Db	419 PYTIRLVBVLSSNGSTSOSAVCGSTLALMDAGVPIKRPAGIAGMLTISEDLSKEAV--	476
Db	514 SVLGLITDVDPSSGEKIDKDYRIVTDIGLEDHGDMDPKLAGTRDGVTAOLDIKPGIPLD	573	QY	510 YRLTDIGIILEDYNGDMDPKINGTKNTGKITALQADIKLPGNIKPIKIVMEIQQASVAKKEIL	569
QY	553 IVMELAIQQASVAKKEIQLQMKITSKPRAKSKENGKPVWETVQVPLSKRAKFVPGGYNIK	612	Db	477 --ITDQGLEDFTGDMDFPKVAGTEK3ITAQYDTHIGLSCYCKTAINDKARLFTL	533
Db	574 IVCESLENARBARQIQLDINERINNSPROGDGASPRALATIKYENDSLRUTLIGPMVLR	633	QY	570 QIMMKTKSPRASRKENGKPVWETVQVPLSKRAKFVPGGYNIKQIQTGVTISQVDBBT	629
QY	613 KLOAETGVITISQVDBETPSVFAPTPSVMHAR--DFTEBICKDQDQEQQLEFGAVYATI	669	Db	534 EKVKACINEPRKLSVYPRAVYNTIDKIRKLGQKTKINKIETGVKIDREGT	593
Db	634 KIEVETGARLIS-IDNGTILITVAKQDMEKAQEQVDFTI-----GRSLVUGVAKGTV	685	QY	630 FSVFAPTPSVMHARDFTEBICKDQDQBOLEFGAVYATITBIRDGTGWNVKLXPNMTAVL	689
QY	670 TEIRDGTGWMYKLYPNNTAVLHNTOLDNE 698		Db	594 VFVLSDDADSANRALKMIDLTKD---VKVGEVLYGKVKTINFGAFVEVLPKGEG-L	647
Db	686 SSIKEYGAFV-E-FPGQQGQLLHMSBLSH 713		QY	690 LHWQDLDNERLN 701	
			Db	648 VHISKLDINKVN 659	
RESULT 2					
B97123	polyribonucleotide nucleotidyltransferase [imported] - Clostridium acetobutylicum				
C;Species:	Clostridium acetobutylicum				
C;Date:	14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001				
C;Accession:	B97123				
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Bacteriol. 183, 4823-4838, 2001	Referece number: A96900; MUID:21359325; PMID:21359325				
A;Title: Genome sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum	A;Accession: B97123				
A;Status: preliminary	A;Molecule type: DNA				
A;Residues: 1-703 <KUR>	A;Cross-references: GB:AE001437; PID:915024781; GSDB:GN00168				
A;Experimental source: Clostridium acetobutylicum ATCC824	A;Gene: CACI808				
C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain					
Query Match 33.9%; Score 1204.5; DB 2; Length 703;					
Best Local Similarity 39.6%; Pred. No. 4.5e-68;					
Matches 266; Conservative 125; Mismatches 258; Indels 23; Gaps 10;					
QY	33 DLGNRKRKEISSLGKLARFADGSAVAVQSGDTAVMVTAWSKTKPSRS-QFMPILVUDYRQAKA	91	Db	889901 EKVKACINEPRKLSVYPRAVYNTIDKIRKLGQKTKINKIETGVKIDREGT	
Db	8 DIAGRKKVKECGKGMLSNAMF1SYGDTVMVNVNASEKPREGIDFFPLSIEYERQVS	67	C;Species:	polyribonucleotide nucleotidyltransferase [imported] - Staphylococcus aureus (strain N31)	
QY	92 AGRIPTNVLREVGTSDEKITSRIDSRTPLRPGYFTDQTLNLAVDGYNEDV	151	C;Date:	Staphylococcus aureus (strain N31) #sequence_revision 10-May-2001 #text_change 22-Oct-2001	
Db	68 VKGKIDPGPVGKEGRBSEKSLHARAIDPRPFLPKGKYRNDVQVCTVMSVEQNTPEIL	127	C;Accession:	C;Species: Staphylococcus aureus	
QY	152 AINGASVALSISDTFWNGPGVAVRIGIDGEYVNVPTRKEMSSSTNLVAGPKSQIVM	211	C;Date:	Staphylococcus aureus (strain N31) #sequence_revision 10-May-2001 #text_change 22-Oct-2001	
Db	128 AMNGASMALCSDIPTTPATVSGCICDFKPVNPLTBEREKSLDLTC-ATNTERVM	186	Db	2 SQEIKVKEFWGRSLTETGQLAQKQANGAVLRYGDTWVLSTASHPRDCDFPFLTV	61
QY	212 LEASABENTIQQDFCHAIVKGKVYQIQIGIQLQVLTGKPTKTPKLSFPTSPETVKYH	271	QY	24 SAGSRAVAVDLGNRKRKEISSLGKLARFADGSAVAVQSGDTAVMVTAWSKTKPSQSOPMPVW	83
Db	187 LEAGADEBIPEDLMLAIDGFGNAODIVQPKAMKEFKGKEKVTPE-LYHKESEIBKDT	245	Db	2 2 SQEIKVKEFWGRSLTETGQLAQKQANGAVLRYGDTWVLSTASHPRDCDFPFLTV	61
QY	272 KLMERLYAVFTDVRHKDVKSRDNEAVKIRIDTEBOLKEKEPPAPAYEIESFNTVAKF	331	QY	84 DYRQAKAAAGRIPTNVLREVGTSDEKITSRIDSRTPLRPGYFTDQTLNLAVD	143
Db	246 EFAFSISIKEIM--YTDRDRERNLRKEKEKISNEFAKIDP-DCADIDBVVYIQLKKV	302	Db	62 NYEERKMYAAGKIPGFKKGEGPGDADATIARLDRPPLPKGKIDVQIMMNLSD	121
QY	332 RSVIUNBYKRCGGRDOLTSLNNSCEVDMFKLHLGSLFQGQTOQLCTWTFDSLESGIK	391	QY	144 GWEPRDVLAVGASVALLSIDIWNGPGVAVRIGIDGEYVNVPTRKEMSSSTNLVAG	203
Db	303 RNMILKEHRRPDRGDRDETRPISCOVDLPLRTHGSLGLPTGLQYNTVTTL--GPID	358	Db	122 PDCSPQMMAMIGMSWALMSDIFQGPTAGVNGYDCKYIINPTVEKRSRDLKVEAG	181
QY	392 DQVITAINGKDKNIMLHYFPPYATNEIGKVTGLNARRELGHLAELAKALYPVIP--RDF	449	QY	257 QKLFTPS-----PEIVKXTHKLMERLYAVFTDVEIDKVSDEAVNKRLDTEBQJK	309
Db			Db	204 APKSQIVMIEASAENILQDQPCRAIKV-----VKTQOIQILOGIQLVKETQVTKTP	256
QY			Db	182 -HKDQAVNWTAEAGASBITHQEMBLAIFGHEHEIQLVDFQOQVYDHIQYKQB-----	232
Db			QY	310 KFPKADP-----YEIISFNVWAKEVRSIVRNEYKRCGDRGDSLTSLNNSCEVDMFKL	363
QY			Db	282 FIDEDPENELIKEVYATLNEVYKEFRLADEKIRPDKGRKEDBIPLDSEBVGILRT	341

QY	364 HGSALLLGGGGQVQLCTWYFDBLESIGKSDVITAAINGIKDKNFMMLYEFPPYATBKGV 423
Db	342 HGSGGFTRGQTOAQLSVITLGAL---GDYQDIDGLCPPEEKRFMHYNFPNFSVGETGPV 397
Qy	424 TGLNRLBLGHGALAKALGKALYVTPP-RDPFTIRVSRVLESGSSMASACCGSALMDS 481
Db	398 RAPGRBREIGHGALGKALYVTPPDTADPYTIRIVSRVLESGSSQASIGSTALMDA 457
Qy	482 GVPISAVAGVAGVAILGKLYTQTDPEKGBIEDBYRLTDLGEDYNGDMDPKIAGTNKITALO 541
Db	458 GVPKAPVAGIATGMLVRED-----SYTIDIOCMEDALGMDPKVAGTKEGITAQ 510
Qy	542 ADIKUGIPIKIUMMAIQQASVAKELIQLQMKTTISKPRASKENGPPVETVQVPLSKRA 601
Db	511 MDIKDGLTREBIALEQARGRGLTINHMLQDPTTELASVAPKVUTMTPDKIR 570
Qy	602 KFVGGGGYNLKKLQQAETGVTISQVDBETFSVAPTPSMHARDFTIECKODQSQLEF 661
Db	571 DVIGGGKKKINBIDTGVKLDIQCCTIFGAVDAMINRARELIEITR---EAEV 625
Qy	662 GAVVYATITEIRDGTGNNVLUKPNTMVTILHNTQDNERL 700
Db	626 GQTYQATVKRICKYGAFAVGLPFGKDA-LLHISQISKNR 663
RESULT 4	
AE2586	polyribonucleotide nucleotidyltransferase pnpa [imported] - Agrobacterium tumefaciens (Eschweiler, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, R.; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science, 294, 2317-2323, 2001)
C;Species: Agrobacterium tumefaciens	
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002	
C;Accession: AE2586	
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.	
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Steier, E.W.	
A;Reference number: AB2577; MUID:21608550; PMID:11743193	
A;Accession: AB2586	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-713 <KUR>	
A;Experimental source: strain C58 (Dupont)	
C;Genetics:	
A;Gene: pnpA	
A;Map position: circular chromosome	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Qy	F97358 polyribonucleotide nucleotidyltransferase (PA4740) [imported] - Agrobacterium tumefaciens (Eschweiler, G.; Gillet, W.; Grant, C.; Guenther, D.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science, 294, 2317-2323, 2001)
C;Species: Agrobacterium tumefaciens	
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002	
C;Accession: F97358	
R;Gooneratne, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldman, A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Goldman, A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Goldman, A.; Reference number: A97359; MUID:21608551; PMID:11743194	
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens	
A;Authors: Gooneratne, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldman, A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Goldman, A.; Reference number: A97359; MUID:21608551; PMID:11743194	
A;Accession: F97358	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-713 <KUR>	
A;Cross-references: GB:AB007869; PIDN:AAK85903.1; PID:915154950; GSPDB: GN00169	
C;Genetics:	
A;Gene: AAK85903.1	
A;Map C:124	
A;Map position: circular chromosome	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Qy	F97358 polyribonucleotide nucleotidyltransferase (PA4740) [imported] - Agrobacterium tumefaciens (Eschweiler, G.; Gillet, W.; Grant, C.; Guenther, D.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science, 294, 2317-2323, 2001)
C;Species: Agrobacterium tumefaciens	
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002	
C;Accession: F97358	
R;Gooneratne, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldman, A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Goldman, A.; Reference number: A97359; MUID:21608551; PMID:11743194	
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens	
A;Authors: Gooneratne, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldman, A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Goldman, A.; Reference number: A97359; MUID:21608551; PMID:11743194	
A;Accession: F97358	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-713 <KUR>	
A;Cross-references: GB:AB007869; PIDN:AAK85903.1; PID:915154950; GSPDB: GN00169	
C;Genetics:	
A;Gene: AAK85903.1	
A;Map C:124	
A;Map position: circular chromosome	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Query Match 32.8%; Score 1166.5; DB 2; Length 713; Best Local Similarity 39.1%; Pred. No. 1.2e-65; Matches 265; Conservative 133; Mismatches 251; Indels 29; Gaps 12; C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain	
Qy	29 AVAVDGLNRKLEISSSKLARPADGSAVQOSDTAVNTAVSKTKPSPSQ-FMLPLVDYRQ 87
Db	6 SVEIWAQGRPLKLETKVARQADGAVIATGVTMVAISAKSPKPGQDFPLTVNQ 65
Qy	88 KAAAGGKPTYLRRGTSKELTSRIDSIRSLRPLPAGFYTQVLCLNLAUDGVNE 147
Db	66 KTYAAGKPTGKFGKGRPRKEKTVLSRSLDRPPLPFGSYKNQDQVWVYVQDVLBLEN 125
Qy	208 QIVMRLASAENILQODPCHA1KVGKVKTQIQLQOLVKETGVTKRTPQQLFTSPETV 267
Db	148 PDVLAINGASVALSISDIPNGGPVAVRIGTIDGEVWNPTRKEMSSSTLNUVAGAPS 207
Qy	185 AVLMVSEAKBLNE1MLGAVMFGKQFQVDAI---IKLAEVAKERB-FEPEDFSA 240
Db	126 PDVLSKVAAHALTSLGIPFMGPVGGARVGYINGEVNLPHDDESDVLDLVAGT-QD 184
Qy	268 KYTHKLAMERLYAVFDYE-HDKVSRDEAVKIRLDEEQLEKEKEPEA--DPYETIESFN 324
Db	241 LNEMLGIAST-ELSTYAKTETKARVAAUAVKTKAFT-PRGKAKT-SPRGKAKT-SPRGKAVF 299
Qy	208 QIVMRLASAENILQODPCHA1KVGKVKTQIQLQOLVKETGVTKRTPQQLFTSPETV 267
Db	148 PDVLAINGASVALSISDIPNGGPVAVRIGTIDGEVWNPTRKEMSSSTLNUVAGAPS 207
Qy	185 AVLMVSEAKBLNE1MLGAVMFGKQFQVDAI---IKLAEVAKERB-FEPEDFSA 240
Db	126 PDVLSKVAAHALTSLGIPFMGPVGGARVGYINGEVNLPHDDESDVLDLVAGT-QD 184
Qy	268 KYTHKLAMERLYAVFDYE-HDKVSRDEAVKIRLDEEQLEKEKEPEA--DPYETIESFN 324

Db 241 LENENMGLAFT-EURTAYKITEKAARYAADVAKTKVKVKAHLFPLPERGEAKYSPERIGAVK 299
 Qy 325 WAKEVFIRSTVLTNEKRCGDRDLTSRNSCCEVDMFKTLGKGSALFORGTOVLTVTEDS 384
 Db 300 HLOQAKIYVRWNLVDTKSRIDGRDLSTVRPIVSEVGVLPRTRGKALPFRGTOVLTATGT 359
 Qy 385 LESGKSDQVITAIANGIKONFMYHYPYATNEIGKUTGLNRELHGKALAKALYV 444
 Db 360 GR---DEOTVDSLTDGMYKERFLPLNPPYVSERGKMSPGREBIGHGKLAIRME 415
 Qy 445 P-RDPFTIRVTSBVLNGSSMASACGGSLALMDSGWPISAVAGVAIGLVTKDP 502
 Db 416 LPTAQOPPYTILRVWSEITNGSSMATVGTSLALMDAGVPLKPVAGIAAGMGLL--- 471
 Qy 503 EKGEDYRLTDIGLIEDTNGDMDFKLAGIPKAGTNGKITALQDILKPGIPKIVMRAIQAS 562
 Db 472 ---EGERPAVSLDILGDEDHGMDPKVAGTAGITSLQMDIKLAGITERIMKTALEOAQ 528
 Qy 563 VAKKEIQLQMKTKSPRASKRKGKNGPWTQVPLSKRACKFGGYNKQKQAEVTGVTI 622
 Db 529 GGRKHLIGMENANATTESSRGOGEPAPRINIPVTKIREVIGSGGVKIREVIGTAKI 588
 Qy 623 SQVDEETFSVFAPTPSVHARDFTIECKDQDQQLERGAVVATITIEIRDGVWKLY 682
 Db 589 NIEDDGTVKIASASGKEIAKRMKWHISIVAAEP---EVQGIVECTVVKTDDFGAFVNFP 643
 Qy 683 PNMTAVLILNTOLDNRL 700
 Db 644 GARDG-LVHISQLASERV 660

RESULT 6

Ado501 poly nucleotide phosphorylase [imported] - *Salmonella enterica* subsp. *enterica* serovar Ty
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 C;Note: this species has also been called *Salmonella typhi*
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AD0901
 R;Parikh, J.; Dougan, G.; James, K.D.; Thompson, N.R.; Pickard, D.; Wain, J.; Churcher, Th.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O Gaoira, P.
 Nature 413, 848-852, 2001.
 Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Ty. Reference number: AB0502; MUID:21534947; PMID:11677608
 Accession: AD0901
 Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-711 <PAR>
 A;Cross-references: GB:AL513382; PIDN: CAD07802.1; PID:gi16504350; GSPDB:GN00176
 C;Genetics:
 A;Gene: STM3463
 C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

Query Match 32.6%; Score 1161; DB 2; Length 711;
 Best Local Similarity 39.9%; Pred. No. 2.6e-65;
 Matches 270; Conservative 118; Mismatches 249; Indels 40; Gaps 11;

Qy 35 GNRKURBISGKLRFADGSAVQSGDTAVMTAVSKTPPSQ_FMLPLVUDYRQKAAG 93
 Db 12 QGHTVTLGETGAMARQTAAYVNSMDTAVTIVQGQDKPFRPLTNYQERTYAG 71
 Qy 94 RIPTNLLRREVGTSDKEBLTTSRITDSSRPPFPGSYFPTQVLCMILAVDGVNEEDVLL 153
 Db 72 RIGPSFPRREGRPSEGTLIARLIDPRVPLPPEGFVNEVQVIAVWSVNPQNPDIVAM 131
 Qy 154 NGASVALSLSPDIPWGPVAVGIGTIDGEVYNPTRKEMSSSTMVWAGPKSQI 213
 Db 132 IGASAMLSLSPCIPNGPPIGAARVYDQVYVNPQDELKSKDULVAGT-EAALMVE 190
 Qy 214 ASAENILQDQPCHAIKVGVKTYQOIQGIOOLVKETGVTKTPQKLFSTSPRIK---- 268
 Db 191 SEAEILSEDNTLGAIVFGHSGQQVVIQIAINDLVKEAGKPRWDWQ-----PEAVNDALNA 244

RESULT 7

Qy 269 YTHKLAMERL---YAVFTDYEDKVSRLDEAVNKRLDTEEQLEKEKFPPEADPVIESEFN 325
 Db 245 RVALAELAERSLSDAYI---TDQKQERIAQDVQIKSEIEQIAEDSFELDANLGEHLHA 299
 Qy 326 WAKEVFIRSTVLTNEKRCGDRDLTSRNSCCEVDMFKTLGKGSALFORGTOVLTVTEDS 385
 Db 300 IERKUVRSVLAGEPRIDGREKOMRIGDVRVTLPRTRGKALPFRGTOVLTATGT- 358
 Qy 386 ESGKSDQVITAIANGIKONFMYHYPYATNEIGKUTGLNRELHGKALAKALYV 445
 Db 359 ---ARDAQVTLDELMGERTDSDPLPHNFYNSPYSCTGWTGSPKREIGGLAKRGVLA 415
 Qy 446 P-RDPFTIRVTSBVLNGSSMASACGGSLALMDSGWPISAVAGVAIGLVTKDP 503
 Db 416 PDMKEPTVTRWSEITNGSSMASCGSLALMDAGVPLKPVAGIAAGMGLL--- 473
 Qy 504 KGEDEYRLTDIGLIEDTNGDMDFKLAGIPKAGTNGKITALQDILKPGIPKIVMRAIQASV 563
 Db 474 ---NYVVLSDILGDEDHGMDPKVAGSRDGSALQMDIKIEGITEIMQVALNQKG 528
 Qy 564 AKKEIQLQMKTKSPRASKRKGKNGPWTQVPLSKRACKFGGYNKQKQAEVTGVTI 623
 Db 529 AKELEIQLQMKTKSPRASKRKGKNGPWTQVPLSKRACKFGGYNKQKQAEVTGVTI 588
 Qy 624 QVDEETFSVFAPTPSVHARDFTIECKDQDQQLERGAVVATITIEIRDGVWKLY 683
 Db 589 NEDDGTVKIASASGKEIAKRMKWHISIVAAEP---EVQGIVECTVVKTDDFGAFVNFP 643
 Qy 684 PNMTAVLILNTOLDNRL 700
 Db 644 GKEG-LVHISQIADKRV 659

RESULT 8

Qy 97782 hypothetical protein pp [imported] - *Rickettsia conorii* (strain Malish 7)
 C;Species: *Rickettsia conorii*
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C;Accession: G97782
 R;Ogata, H.; Audic, S.; Renesto-Audifren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Roig, H.; 293, 2093-2098, 2001.
 A;Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.
 A;Reference number: A97700; MUID:21442074; PMID:1557893
 A;Accession: G97782
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-749 <PAR>
 A;Cross-references: GB:AE006914; PIDN:AA03201.1; PID:gi15619751; GSPDB:GN00173
 C;Genetics:
 A;Gene: pnp
 C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

Query Match 32.6%; Score 1158.5; DB 2; Length 749;
 Best Local Similarity 38.0%; Pred. No. 4e-65;
 Matches 257; Conservative 145; Mismatches 244; Indels 31; Gaps 12;

Qy 31 AVDGNRKRURBISGKLRFADGSAVQSGDTAVMTAVSKTPPSQ_FMLPLVUDYRQKA 89
 Db 8 SVTNGQVLELSTKIAQDAGTVKVGNSVILCTAVWANKAKEGICGFLPLTNYRME 67
 Qy 90 AAAGRIPTYLRLREVGTSDKEBLTTSRITDSSRPPFPGSYFPTQVLCMILAVDGVNEEDVLL 149
 Db 68 YAACKIPSGCPKHKGKAQDREVLSRUDPRPLFRAFVNTHVTCVSLSYDPTEDV 127
 Qy 150 VLAINGASVALSLSPDIPWGPVAVGIGTIDGEVYNPTRKEMSSSTMVWAGPKSQI 209
 Db 128 ILAIGASAMLSLSPCIPNGPPIGAARVYDQVYVNPQDELKSKDULVAGT-EAALMVE 186
 Qy 210 VMLRASMTLQDPCHAIKVGVKTYQOIQGIOOLVKETGVTKTPQKLFSTSPRIK----- 269
 Db 187 MMVSEAHILSEEDMLAIVKGFGBESFQVQIKITKELABEAKKPKLEMOPALYPS-LIGK 244

QY 270 THQL---AMERLAVFTYEDKVSRSRDAVNRKURLDTEBQLKEKFPEDPYRIESENWV 326
 Db 245 IEKLFVKIEQFAIKSIQER- STNLDLIPPEKVLTHFVSDIENK--KYSNQIRESKAI 301
 QY 327 AKVFRSVLNEKRCRDLSRSLRNVSCEVDMRKTLLHGSALFORQTLQVLCVTF-DSL 385
 Db 302 ESDILRNILEKURRIDRSTTDIROIACIGLPSAHSALTRGETQSLVSTTFCSL 361
 QY 386 ESGIKSDQVITATNGIKKNFMAYEFPYPATNEIGKVTGLARRELGHGALAKALYPI 445
 Db 362 D----EQIVDSLAEYEKFERMAYIIFPPSYNEAMPKAPSRRVGHGKLAIRNPL 416
 QY 446 PR--DFPPTIRVSELENSGESSMASAGGSLALMDSGGVPISSAVAGVAGLUTKDPB 503
 Db 417 PNKVQFPPISRVVAAETTESNGGSSMATVCGSSLALMAYVPIKAVAGIANGLVK--- 471
 QY 504 KGIEDYRLLTDIGLIEYNGDMDPKLAGTNKITALOADIKLPGITKIVMELAQASV 563
 Db 472 --EGKINPAVLSLDIGDDEDYFGMDPKVAGTSSGITALQMDIKSGUDFKIMVKAEGRL 529
 QY 564 AKKEILQIMKNTISKPRASKRKENGPVVEVTPQVLSPKRASKRKFVPGGGYNUKKQAEATGVTIS 623
 Db 530 GRHLILEOMKVKISKPNELSKAPSTTIKIDKDKLDRDGGKIKLICETSGAKID 589
 QY 624 QVDEETTSVFAFPSPVMEHARDTTECKDQBQLQFLGAVVTTATBIRDGVMVLYP 683
 Db 590 ISDGTIVSVYASDRDKLKVADLKIKATVVEP----BIEGFNTVVKVULDSGAFINTVG 644
 QY 684 NMVAVLHLNTQLDNERL 700
 Db 645 NKQG-FVHISEVSGERI 660

RESULT 8

E9134
 Polyribonucleotide phosphorylase [imported] - Escherichia coli (strain 0157:H7, substrate R1)
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C;Accession: E9134
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iibai, K.; Yokoyama, K.; Han, C.G.;
 gasanara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-26, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and 96
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: E9134
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-734 <HAX>
 A;Cross-references: GB:BA000007; PID:91336518; GSPDB:GN00154
 A;Experimental source: strain 0157:H7, substrate R1D; 0509952
 C;Genetics:
 A;Gene: BC4045
 C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

.. Query Match 32.6%; Score 1158; DB 2; Length 734;
 .. Best Local Similarity 40.0%; Pred. No. 4.2e-65; Matches 270; Conservative 117; Mismatches 252; Indels 36; Gaps 11;

QY 35 GNRKLISSGKLRPADGSAVQVQSGDTAVMTAVSKTKPSQ-FMPLVVDYRQKAAAG 93
 Db 35 GQHTVLTGMMARQATAVWMSMDPAVFTVVGQKAKCQDFPFLTVQYQERTYAG 94
 QY 94 RIPTNVLREVGTSKELTSRITRSTRPLPAGYFYDVTQVLCNLAVDGYNEDVLA 153
 Db 95 RIGSPFRRGREGSEGEETLARLIDRIPRLPPEGFNEQVIATVSVNQNPNDIVAM 154
 QY 154 NGASVALSLSDIPWNPQGAVGIGIDGEGYVNPTRKMSSTLNVLVAGPKSQIVMLE 213
 Db 155 IGSAAALSLSPFNGPPIGARVGYINDQVYVQPTODELKSKLUDLVAGT-EAVLME 213
 QY 214 ASAENTIQQDFCHAIKGVKVTQOIQIOQVJKEVTKTPQKLFPTSP--EIVKYT 270
 Db 214 SRAELSLSDQMLGAVVFGHEOCQVQVIONINELVKEAG---KPRWDWQPEPVNEALNARV 269

RESULT 9

H85979
 Polyribonucleotide phosphorylase [imported] - Escherichia coli (strain 0157:H7, substrate ED)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
 C;Accession: H85979
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.J.; Grobecker, E.J.; Davis, N.W.; Lim, A.; Dimarzio, E.; Potamitis, K.; Apodaca, Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: H85979
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-734 <STO>
 A;Cross-references: GB:BA005174; PID:91251775; PID:AGG58300.1; GSPDB:GN00145; UWGP:Z45
 A;Experimental source: strain 0157:H7, substrate EDL933
 C;Genetics:
 A;Gene: pop
 C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

.. Query Match 32.6%; Score 1158; DB 2; Length 734;
 .. Best Local Similarity 40.0%; Pred. No. 4.2e-65; Matches 270; Conservative 117; Mismatches 252; Indels 36; Gaps 11;

QY 35 GNRKLISSGKLRPADGSAVQVQSGDTAVMTAVSKTKPSQ-FMPLVVDYRQKAAAG 93
 Db 35 GQHTVLTGMMARQATAVWMSMDPAVFTVVGQKAKCQDFPFLTVQYQERTYAG 94
 QY 94 RIPTNVLREVGTSKELTSRITRSTRPLPAGYFYDVTQVLCNLAVDGYNEDVLA 153
 Db 95 RIGSPFRRGREGSEGEETLARLIDRIPRLPPEGFNEQVIATVSVNQNPNDIVAM 154
 QY 154 NGASVALSLSDIPWNPQGAVGIGIDGEGYVNPTRKMSSTLNVLVAGPKSQIVMLE 213
 Db 155 IGSAAALSLSPFNGPPIGARVGYINDQVYVQPTODELKSKLUDLVAGT-EAVLME 213
 QY 214 ASAENTIQQDFCHAIKGVKVTQOIQIOQVJKEVTKTPQKLFPTSP--EIVKYT 270

QY	214 ASAENIQQPDECHATKUGVKVTKQTQGIOOLVKENG-----VTKRTPOKLFPSPEV	267
Db	220 CQAGTQVSEDDVLGVATEFAHABMOGVYALIBQRAEVGHEKENFLABEGPANDYV-----	274
Db	270 AALARBLSDAYR-----DQKQEEVAYQDVQVKISTATLAEDTLDENELGEHLHAE 324	
QY	275 ELTEKAKAAGJRDALLTHGKDSRSRSLKALRNLGLEYGVYDPTASGSABLTQA---LKD 330	
Db	277 KTHKLAMERI-KVFTDYEDKVKRSDEAVNKIRL-----DTEBOLKEKEFPEADPYEIE 321	
QY	278 KEVRSTIVLNVEYKRGDRDLSLRNVSCEVMFKTLHGSALFORQTOVQTVTDSLES 387	
Db	279 KNNVYRSVLAGEPRIDGKREKDMIRGDLDRVTLGVLPLRTHGSALFTRQTOVATGTT----- 381	
QY	388 GIKDQVITTAINGKDKNEMLYHEPPYATNEIGVTLGRLRELGHALAKALYVPI- 446	
Db	382 -ARDAQVLDLBMGERDTDFLPHYNPPYSPVGETGNGSPKRREIGHGLAKRGVLAQMPD 440	
QY	447 -RDFPFTIRTVSEVLENSGSSMASACGSALMDSGVPTSSAVAGVALGLVTKDPEKG 505	
Db	441 MDKFPFTVRVVSETESSNGSSMASVGASLALMDQGVPIKAVGJAMGLVKGGD--- 496	
QY	497 --NYVVLSDLGDLGDDHLDGMDFKVAGSRDGSIALQMDIKIEGTEIMVALNQAKGAR 553	
Db	506 EIBYRLLDUDLGLDHYNGMDPKAGTNKGITAQDIKLPGIPKIVMEIAQDASVAK 565	
QY	554 KEIQIOMNKTKISKPRSKRKGKNGPVETVQPLSKRDKKFKVOPGGNMLKKLOETGTVISQV 625	
Db	555 LHTIGVMEQAINARPGDSEFAPRHTIKINPKDKDVKVKGGSVIRALJETGTEIE 613	
QY	626 DEETFSVFAFPTPSVHEDRPFTECKDQEQQLFEGAVVATATTEIRDGVMTVLYPNM 685	
Db	614 DDGTVKIAATDGEKAKHAIRRIEE-----TAEIIVGVRVYTKRIVDGFAGVVAIGGG 668	
QY	686 TAVLHNTQLDNERL 700	
Db	669 EG-LVHISQIADKRV 682	
RESULT 10		
QY	G75320 polynucleotide phosphorylase - <i>Deltaproteobacterium</i> radiodurans (strain R1)	
C.Species	Deinococcus radiodurans	
C.Date	03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000	
C.Accession	G75320	
R.White.O.	Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Zalewski, C.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma	
S.	M.; Sheen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma	
S.	Science 286, 1571-1577, 1999	
A.Title	Genome sequence of the radioresistant bacterium <i>Deinococcus radiodurans</i> R1.	
A.Reference number	AY75250; MUID:20036896; PMID:10567266	
A.Accession	G75320	
A.Status	preliminary	
A.Molecule type	DNA	
A.Residues	1-810 <WHI>	
A.Cross-references	GB:AB002042; GB:AB000513; NID:96459848; PIDN:AAF11608.1; PID:9645985	
A.Experimental source	strain R1	
A.Genetics		
A.Gene	DR2063	
A.Map position	1	
C.Superfamily	polyribonucleotide nucleotidyltransferase alpha chain	
Query Match	32.5%: Score 1156; DB 2; Length 810;	
Best Local Similarity	38.1%; Pred. No. 6.5e-65; Matches 261; Conservative 138; Mismatches 236; Indels 50; Gaps 14;	
QY	34 LGNRKLEISSGKLLRFAFGDSAVVQSGDTAVMTAVSKTKPSQSPQMPFLVYDQRQAAAG 93	
Db	41 LGGRBLISIENGKLLAKLVSSTVTRGPTLVLVQASDQSKLPLPPLTVEFERRHYAVG 100	
QY	94 RIPTYLRIRVEGTSDEKILSRIDRSTRPLFFPAGYFYDPTQVQLCMLLAVDGVNEDBDLAI 153	
Db	101 KIPGSPFORRERGRGPGEKALSARITDQIRPLPKPKYRHEQIVITVLSADQGNAQBDVLP 160	
QY	154 NGASVALSISDIPWPGPVGAVRIGIDGEVWVNPTRKENSSTLNLVWQGPKSQIVMLE 213	
Db	161 IGAAALASISDIPWPGPCTAVRSGIDQGTVVNPTEOLRSRMDLWVAGT-REMMVME 219	
RESULT 11		
QY	S74509 polynucleotide nucleotidyltransferase (EC 2.7.7.8) alpha chain - <i>Synechocystis</i> sp.	
N;Alternate names	protein SII1043	
C.Species	<i>Synechocystis</i> sp.	
A.Variety	PCC 6803	
C.Date	25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000	
C.Accession	S74509	
R.Kaneko	T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Yamada, M.; Yabuuchi, O.; Okumura, S.; Shimojo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yabuuchi, DNA Res. 3, 109-136, 1996	
A.Title	Sequence analysis of the genome of the unicellular cyanobacterium <i>Synechocystis</i> sp.	
A.Reference number	S74322; MUID:97061201; PMID:8905231	
A.Accession	S74509	
A.Status	nucleic acid sequence not shown; translation not shown	
A.Molecule type	DNA	
A.Residues	1-718 <KAN>	
A.Cross-references	EMBL:D90899; GB:AB001339; NID:91651650; PIDN:BAA16661.1; PID:g16517	
A;Note	the nucleotide sequence was submitted to the EMBL Data Library, June 1996	
C.Genetics		
A;Gene	ppn	
C.Superfamily	polyribonucleotide nucleotidyltransferase alpha chain	
C;Keywords	nucleotidyltransferase	
Query Match	32.4%: Score 1153; DB 2; Length 718;	
Best Local Similarity	37.5%; Pred. No. 8.3e-65; Matches 257; Conservative 131; Mismatches 261; Indels 36; Gaps 10;	
QY	31 AVDGRKLEISSGKLLRFAFGDSAVVQSGDTAVMTAVSKTKPSQSPQMPFLVYDQRQAA 90	
Db	7 SISPDQDIRKMLTAPQAGGSVLIQDPAVLTATRAGKRGDGTDFLPLTVDVYEGRLY 66	
QY	91 AGRPTYLRIRVEGTSDEKILSRIDRSTRPLFFPAGYFYDPTQVQLCMLLAVDGVNEDBDLAI 150	

Db	67	AAGRIPGFLREGPRPEKATLISLNUIDRPRPLPHMKAELQVATTLSMDEEVPPDV	125
A;Reference number:	A23984		
A;Accession:	BA3984		
A;Status: preliminary			
A;Molecule type:	DNA		
A;Residues:	'W',25-85 <BWA>		
A;Cross-references:	GB:MA4425; NID:9147747; PID:AAA24596.1; PID:9147749		
A;Note: the authors translated the codon GCC for residue 74 as Arg			
A;Genes:			
A;Map position:	69 min		
A;Start codon:	TTC		
C;Superfamily:	polyribonucleotide nucleotidyltransferase		
C;Keywords:	nucleotidyltransferase		
F;25-734/Product: polyribonucleotide nucleotidyltransferase alpha chain #status predict			
Query Match	32.4%; Score 1152; DB 2; Length 734;		
Best Local Similarity	39.9%; Pred. No. 1e-64;		
Matches	269; Conservative 118; Mismatches 252; Indels 36; Gaps 11;		
QY	319 IIESFFNVVAKEVFERSIVLNVEYKRCGDRDLTSLRNVNSCEVDMF-KTLHGSALFORQQTQVL 377		
Db	302 VGNLYKALTRKLMRQIVDVGVRVGRKLBQVRPSCIEVGLPRLPRVHGSLPFRNRCQTLQVL 361		
QY	242 EFTIARASKKLTITLGQFOLGKDRGDAALDEKATEVETVIAELEBTDPYKQSVEDPKL 301		
Db	186 MVEAGANOLPPEQDTEAIDPGEYAVODLNAOREALMTDGLTATSE---PPRNATE 241		
QY	268 KYTHKLAMERLYAVFTDYEUDKVSDEAVNKIRLDEEOLKEKEFPEADPVE----- 318		
Db	211 MLERAAENLTLQDPCFHAIKVGKVKYQIQIQQVKEGTGVTKRPQKLFITPS---EIV 267		
QY	242 EFTIARASKKLTITLGQFOLGKDRGDAALDEKATEVETVIAELEBTDPYKQSVEDPKL 301		
Db	186 MVEAGANOLPPEQDTEAIDPGEYAVODLNAOREALMTDGLTATSE---PPRNATE 241		
QY	362 SLAT---LSPGDAQDADLHPSPEKYLHVNHYPPVSGEARNRSPCRREIGHGALA 418		
Db	438 EKALYKPVIP---RDPFTTIRVSEVESEVSENGSSMASACCGSIALMDSGVPISSAVAGVAIG 495		
QY	419 BRAIPVLPQEDPFPYVVRVSEIETNSNGSTSMWSGVCGSTALAMDAAGVPKKPYVSGAAMG 478		
Db	496 LVTKDPKEKEIEPDWRLTDLIGDTEYNGDMDFKTAGTNGKITAQADTKLPGPIPKIVM 555		
QY	479 LIKEGD-----EIRLTDIQLBDPLGMDPKVAGTDSGITALQMDMKDGLSMEVVS 531		
QY	556 BAQOQASVAKKEIQLQIOMKNTISKPRASKRKGNGPVVETVQVPLSKRAKFKVPGGGNLUKKLQ 615		
Db	532 KAIQMPALPARLHIDKMLATIREPRPELSPFAPRULIKPEHIGMVIQPGKTKIGIT 591		
QY	616 AETGTVTISQDEETTSVAPFTPSVHEDRFITETCKDQDQEQLRFGAVYATTEBIRD 675		
Db	592 BQTSCKKIDIAADDGTVTIASSSEGGERERAROMYNNTR----KLNKEGVITGRVIRIPI 646		
QY	676 GWMYKLYPNTAVLHNTQNLNERL 700		
Db	647 GAFVFEVLPGKEG-MHHSQLTGEGV 670		
RESULT 12			
H65106	polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) alpha chain - Escherichia coli (E. coli K-12) #text_change 01-Mar-2002		
N;Alternate names: polynucleotide phosphorylase			
C;Species: Escherichia coli			
C;Date: 12-Sep-1997 #sequence revision 23-Jan-1998 #text_change 01-Mar-2002			
C;Accession: H65106; A26118; B23984			
C;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vega, L.A.; Rose, D.J.; Mau, B.; Shao, Y.			
C;Science 277, 1453-1462, 1997			
A;Title: The complete genome sequence of Escherichia coli K-12.			
A;Reference number: A67220; NUID:97426617; PMID:9278503			
A;Accession: H65106			
A;Status: nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-734 <BLAT>			
A;Cross-references: GB:AB000397; GB:U00096; NID:92367199; PIDN: AAC76198.1; PID:91789555;			
A;Experimental source: strain K-12, substrain MG1655			
R;Regnier, P.; Grunberg-Manago, M.; Portier, C.			
J. Biol. Chem. 262, 63-68, 1987			
A;Title: Nucleotide sequence of the pap gene of Escherichia coli encoding polynucleotide in S1.			
A;Reference number: A26118; NUID:87083499; PMID:2432069			
A;Molecule type: DNA			
A;Residues: 'W',25-379, 'R',381-472, 'S',474-734 <REG>			
A;Cross-references: EMBL:J02638; NID:9147744; PID:AAA83905.1; PID:9551833			
A;Title: Promoter activity and transcript mapping in the regulatory region for genes encoding			
RESULT 13			
AC3497	polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) [imported] - <i>Brucella melitensis</i>		
C;Species: <i>Brucella melitensis</i>			
C;Date: 01-Feb-2002 #sequence revision 01-Feb-2002 #text_change 15-Feb-2002			

C;Accession: AC3497
 R;DelVecchio, C.; Kaprattal, V.; Redkar, R.J.; Patra, G.; Mulier, C.; Log, T.; Ivanova, C.;Species: *Yersinia pestis*
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AC0424
 C;Accession: AC3497
 A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A;Accession: AC0252; PMID:1175688
 A;Accession: AC0424; PMID:1175688
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-714 <KUR>
 A;Experimental source: GB:AB008917; PIDN: AAL53142.1; PID:gi17984011; GSPDB:GN00190
 C;Genetics:
 A;Gene: BMEII_961
 A;Position: 1
 C;Superfamily: nucleotidyltransferase
 C;Keywords: nucleotidyltransferase

Query Match 32.3%; Score 1150; DB 2; Length 714;
 Best Local Similarity 39.0%; Pred. No. 1.3e-64; Indels 48; Gaps 14;
 Matches 268; Conservative 124; Mismatches 247; Indels 48; Gaps 14;

QY 30 WAVDVLGNRKLEISSGKLRPADGSAWVQSGDTAVANTAVSKTKSPSQ-FMPLVUDYROK 88
 Db 7 VEIENGKGRPITLETSKLRQDGAVLATVUSAKEPREGQDPFPPLTVNQEK 66
 QY 89 AAAAGCRIPTNTRREVGTSDRKILTSRITDRSTRPLFPACYFPYPROVLNLLAVTDGVNBP 148
 Db 67 TYAAGKIPGGYKREGRPSENBTLYSLRDRIPRLPFLVGDYKNDIQVQVTVLQHDLNPP 126
 QY 149 DVLTAINGASVYALSLSDIPWNGPVGAVRIGIDGEVYVNPTRKENSSTSNTLVVAGAPKQ 208
 Db 127 DILSMVAAASALNTISGVFMGPISGARVGYDGEVYLNPNIDEMPESKUDLIVVAGSEA- 185
 QY 209 IVMLESAEMLQDQFCHAIVKGKVYQOIQIOOLVKETGVTKTPQKLUFTSPPEIVK 268
 Db 186 VLMVSESEAQSELPEDMLGATMFGHKSQFYDIAI---IKLAEVAKERPDF---QPEDIS 239
 QY 269 YTHKLMERLYAVFTDYEH---DKVSRDDEBANKIRLDRTEBOLKEKEFPR---AD 315
 Db 240 --ELEAKVLLWENDLREAKYKITEKQARYAIVADAKAKAEEH---FPGEVETMSAE 293
 QY 316 PYELLESFNTVAKEYFVSYIENEYKRCGSDLTSRNVSCEVDEKFTLGSALFORGOTO 375
 Db 294 OPATI---FKHQAKTWRNITDGNRIGDRDLSTVRPVISEVGILPRTHGSAFLPFGETO 351
 QY 376 WLCTTFDSLESGIKSQDVITAINGKDKOFLHYEPFPYATNEIGKVTLNRLBHGAA 435
 Db 352 AIVVATLGDBB---DEOMIDALTCTYKESFMLHNFPYPSVGETGRMGSPGRBIGHCK 407
 QY 436 LAEKALYFVIB---RDPFTTRVTSEVLESLNGSSMSACCGGLALMDSGVPISSAVAGA 493
 Db 408 LAWRATHPMRPLAEQFPYITRAVSEITESNSKSSMATAVCGTSIALMDAGVPIVRVAGIA 467
 QY 494 IGLVLTIDPBERGBEISDYLTDIIGEDYNGMDKIAKGNTKGITALQADIKLUGIPIKI 553
 Db 468 MGLIK-----EGERFAVLSIDLGDDEDHLDGMDFKVGAVTGFITSLQMDKIDGTEI 520
 QY 554 WMEAQOQASVAKKEIQLQIMKTTISPRASKMKGNGVETVQVPLSKRAKEVGGYNNLKK 613
 Db 521 MKVALBOAKSGRVRHILGEMAKAIISSRAELEGEFABRIEVNINTPKIRDVIGSGKVIR 580
 QY 614 IQAEQFTISQVDEETFSVFTAPTPSMWHEARDFTIECKODQEQOLEFGAVYVATITEIR 673
 Db 581 IVEKTKAKINIEEDGIVKVISSNGKEIEAKWVWISIVAEP---EVGBIYEGTVKTA 635
 QY 674 DTGVWVVKLYPNTAVILHNTQDNLRL 700
 Db 636 DFGAFVNFFGRDQ-LVHISQIADRV 661

Query Match 32.2%; Score 1146; DB 2; Length 705;
 Best Local Similarity 40.0%; Pred. No. 2.2e-64; Indels 32; Gaps 10;
 Matches 269; Conservative 113; Mismatches 259; Indels 32; Gaps 10;

QY 35 GNRKLEISSGKLRPADGSAWVQSGDTAVANTAVSKTKSPSQ-FMPLVUDYROKAMAG 93
 Db 12 GOHTVITFGMMARQATAVVMWSMDTAVFTVUFGQKAKPGSOSRFPPLTVNQERTVYAG 71
 QY 94 RIPTVYLARREVGTSDKETLTSRITDRSTRPLFPACYFPYDFTQVCLNLLAVDGWNEPDVIA 153
 Db 72 RIPESPFRREGRESEGEFTSRIDPRPRLPDSFNUEVQVIATVSVNPQNPDIW 131
 QY 154 NGASVALSLSDIPWNGPVGAVRIGIDGEVYVNPTRKENSSTSNTLVVAGAPKQSVIIL 213
 Db 132 IGSASALSLSGIPNGPPIGARVGFINDQVYVNPITDILKESRDLVWAGT-AGAVLNE 190
 QY 214 ASARNILQDPCHAIKVSYKTYQOIQSOLQMKETGTYKTPQKLFIPSPEIVKTH- 271
 Db 191 SEADILSEBOMLGAUVFGHEQQVVIENNALVAEAGPKWDQ---AEPVNEALHAR 245
 QY 272 --KLAMERLYAVFTDYEHDKVSRDDEBANKIRLDRTEBOLKEKEFPEADPYELLESFNVKAE 329
 Db 246 VAELEARIGDADYRITE---KQERTYQDIAKADYTEAALQADDTDAEIQDITLASVEKN 303
 QY 330 VFRSIVLNBYKRCGDRDITSLRNVCEVDMKFTIHSALFQRGQTQVCTVTDLSSEGI 389
 Db 304 WVRSRVLRGEPRIDGREKDMIRGLDVRGILPRLPRTGSAFLTRGETQALVATATGARLAQ 363
 QY 390 KSDQDVITAINGKDKOFLHYEPFPYATNEIGKVTLNRLBHGAAKALYFVIP---R 447
 Db 364 NIDEML---GERTDISFLHYNFPYPCVGETGMGSPKRREICHGRKLVAVMPAS 419
 QY 448 DPFTTIRVTSEVLESLNGSSMSACCGGLALMDSGVPISSAVAGAIVGLVTKDPEKEI 507
 Db 420 EFPYFTIRVVSBEITESNGSSMSAVGASLALMDAGVPIKAAGVAGIAMIQCLVKEGD----- 473
 QY 508 EDYRLIDTIDGIEDYNGMDKIAKGNTKGITALQADIKLUGIPIKIVMVAIIOQASVAKKE 567
 Db 474 -NFWVLSIDLGDDEDHLDGMDFKVGASRQVTAQMDKIEGTRIEMQVALNQAGARLH 532
 QY 568 ILOTMKNTSKPRASKRKGNGPVVETVQVPLSKRAKEVGGYNNLKKQAGETGTTISODRE 627
 Db 533 ILGVMEQATSTPRGDISPAPRIFTYTMKINPEKTDVICKGGSVIRALTDETGTTIEIDD 592
 QY 628 ETEFVAPPFSPSMWHEARDFTIECKODQEQOLEFGAVYVATITEIR 687
 Db 593 GTKIAAATGDKAKHAIIRIEI---TAEIENGRIAGKVTRIVDGFAGFAVAGGGREG 647
 QY 688 VLLIATQDNERL 700
 Db 648 -LVHISQIADRV 659

RESULT 15
G83950
polynucleotide phosphorylase *ppgA* [imported] - *Bacillus halodurans* (strain C-125)

C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: G83950

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A;Reference number: A83600; MUID:20512582; PMID:11058132

A;Accession: G83950

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-704 <STO>

A;Cross-references: GB:AP001515; GB:BA000004; NID:910174886; PIDN:BA06126.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: *ppgA*

C;Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

Query Match 32.2%; Score 1145; DB 2; Length 704;

Best Local Similarity 39.3%; Pred. No. 2.6e-64; Mismatches 251; Gaps 11;

Matches 265; Conservative 132; Indels 26; Gaps 11;

Qy 31 AVDLGKRLKLESSKGKLRFAOGSAVQSGDPAVMTAVSKTPPSQFMPFLVVDTRQAA 90

Db 9 SIDWAGRKLTVENGQQLAKQANGAVLVRQYDIAVLSTASKPKDLPFFPLVNEYERLY 68

Qy 91 AAGRIPTYLRLREVGTSDKEILTSRIDSIRPLPAGYFFDTQVLCLNLLAVDGYNEDV 150

Db 69 AAGKIPGFGFIRKREGPSKEAIIASRLIDPRPLPFFBGFRRREVQVISVMSVQDCSEM 128

Qy 151 LAINGASVALLSIDIPWNGPVGAVRIGIDSEYYVNPNTRKMSSTANLVAGAPKSOIV 210

Db 129 AAWVGSSLLASISDIPFEGPAGAVTWRQIDQFVINPQTQDLEKSDIHLVWAGT-KDAIN 187

Qy 211 MLEASARNITLQQDFCHAIKVSKYKTOIIQG1QQLKETGTGKRTPOKLFTPSPEIVKT 270

Db 188 MVEAGABEVPEVDYEMLEAIFGHEKTRKLEFORKIAEFGKSK-TDVLVKDPMLEQEV 246

Qy 271 HKGLAMERL-YAVFTDYEHDKUSRDAVWKIRLDEEOLKEKEPREADPYHIESFNVAKE 329

Db 247 RYKAESDILKQAVQVPERKHDIAEAMDKVUYETSD-NEVPLSEVNIRHK--IVKE 302

Qy 330 VFRSIVLNEVYKRCGDRQTSLGRNSCEVDMKTYLHSALFQGOTQVLCWTFDSLESGI 389

Db 303 EVRLLIVKEKURPDRGEBIDEKPLSSOVOGILPRTHTGSGLFRQGQALSTCITLGAL--- 358

Qy 390 XSDQVITAINGKDKKOMFLHTEFPPVATNEKGKVUNRRLGHGALAKLYPVPR 447

Db 359 GDVQILDGGIGLGEESKRFRMHMHNPQFQSGETCPIRGPGRRTIGHGALGERALEPVPSEQ 418

Qy 448 DPFPPTIRVTSEVLESLNGSSMASACGSLALMDSGPPISSAVAGVAIGLVVTKTDPERKGEI 507

Db 419 DPPPYTIRLVLSEVLESLNGSTSSACGSLALMDSGPPIKAVAGVAMGLVQD----- 471

Qy 508 EDYRLJTDILG3TEDYNGDMDFKLAGTNGKTTALQADIKLPGSIPKIKVMEAIQOASVAKKE 567

Db 472 EIVSVLTDIQQMEDALGDMDFKVKAGTRKGVTAQMDIKISGIDRAILEQALBQARKGRMI 531

Qy 568 ILOIMNKTISKEKASRKGENGIVYETVQVPLSKRAKTVGPGSYNKLQAEFGVTSQVDE 627

Db 532 IINDNMLAISSESSSELSPYAPKLTWTINPKDIRDVIGPSKOMKNTIEDIGVKDIEQD 591

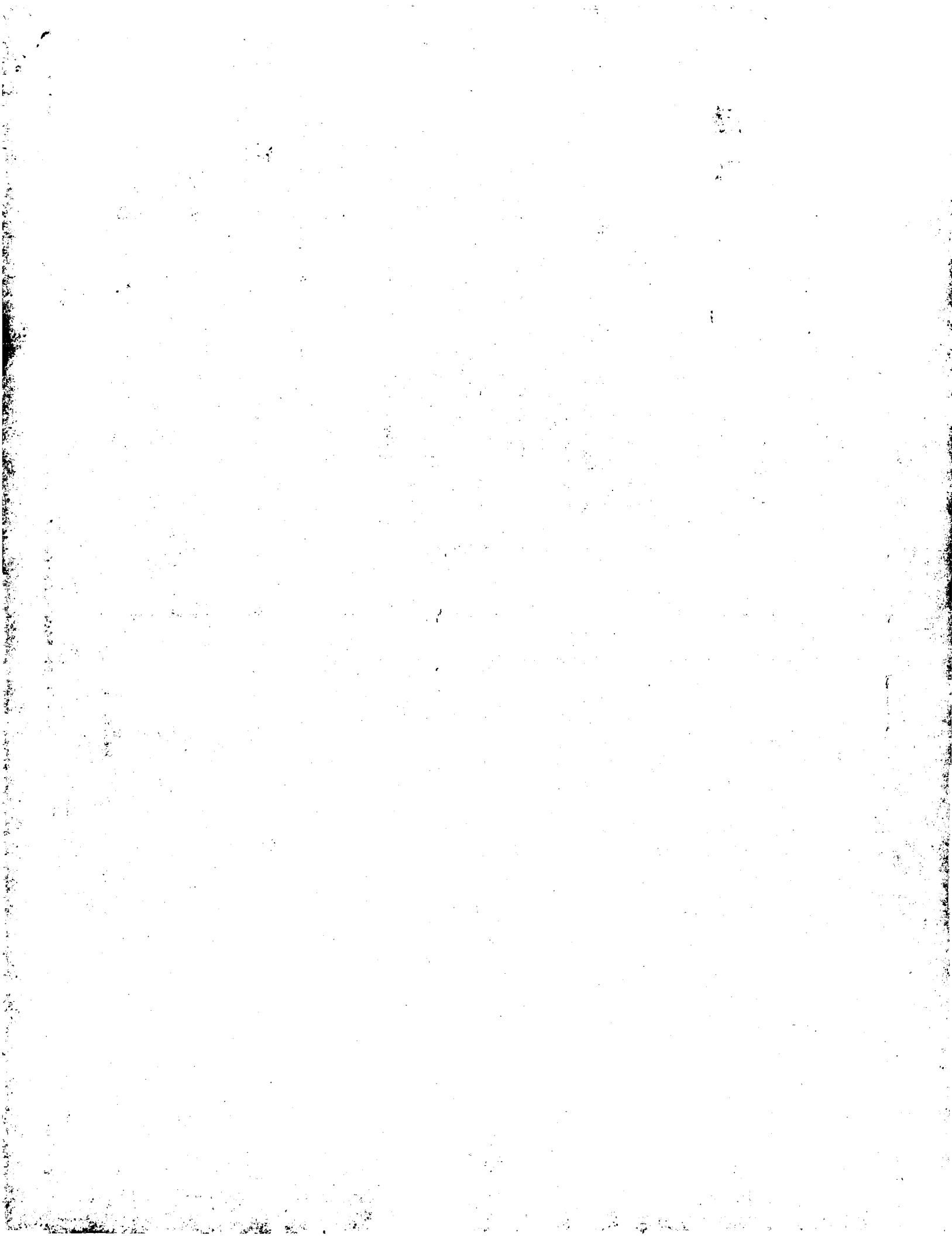
Qy 628 EITFSVFAPTPSMHEARDPTEICKDQEQLEQLEPGAVVATITIEIRDTGWNKLYPNTA 687

Db 592 GPIVYISSADTMNKAKAREIEEDIVR----EVEVGOMYLGTWKRIBKFGAFVLFKGKD 646

Qy 688 VILHNTOLDNERLN 701

Db 647 -LVHISQLAERVN 659

Search completed: January 8, 2004, 10:40:52
Job time : 29 secs



STRUCTURE BY NMR OF 617-692.
 RX MEDLINE:97160844; PubMed=900164;
 RA Bucroff M.; Hubbard T.J.; Proctor M.; Freund S.M.; Murzin A.G.;
 RT "The solution structure of the S1 RNA binding domain: a member of an
 ancient nucleic acid-binding fold";
 RL Cell 88:233-242 (1997).
 CC -1- FUNCTION: Involved in RNA degradation. Hydrolyzes single-stranded
 polyribonucleotides progressively in the 3' to 5' direction. Involved in RNA processing and messenger RNA degradation. Usage by
 -1- CATALYTIC ACTIVITY: {RNA} (N+1) + phosphate = {RNA} (N) + a
 CC nucleoside diphosphate.
 -1- SUBUNIT: Homotrimer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- INDUCTION: In response to low temperature.
 CC -1- SIMILARITY: Contains 1 KH domain.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@ib-sib.ch).
 CC
 DR EMBL: J02638; AAA83905_1; -;
 DR EMBL: U88997; AAA57967_1; ALT_INIT.
 DR EMBL: AB003917; AAC76198_1; ALT_INIT.
 DR EMBL: X00761; CA425332_1; -;
 DR EMBL: M14425; AA424596_1; -;
 DR PDB: 1SMQ; 01-APR-97.
 DR SWISS-2DPAGE; P05055; COL1.
 DR ECO2DBASE; C0880; 6TH_EDITION.
 DR ECO2DBASE; C0880; 6TH_EDITION.
 DR EMBL: M14425; AA424596_1; -;
 DR PDB: 1H61; H65106.
 DR PDB: 1SMQ; 01-APR-97.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR InterPro: IPR00110; Ribosomal_S1.
 DR InterPro: IPR003029; S1.
 DR Pfam: PF00013; KH_1.
 DR Pfam: PF03726; PNAbase_1.
 DR Pfam: PF01138; RNase_PH_2.
 DR Pfam: PF03725; RNase_PH_C_2.
 DR PRINTS: PRO0621; RIBOSOMAL_S1.
 DR SMART: SM00322; KH_1.
 DR PROSITE: PS50084; KR_TYPE_1; 1.
 KW Transferase; Nucleotidyltransferase; RNA-binding; 3D-structure;
 KW Complete proteome.
 FT DOMAIN 557 591 KH.
 FT DOMAIN 622 690 S1 MOTIF.
 FT CONFLICT 357 357 G -> R (IN REF. 1).
 FT CONFLICT 450 450 L -> S (IN REF. 1).
 FT TURN 621 622
 FT STRAND 624 633
 FT TURN 634 635
 FT STRAND 636 640
 FT TURN 651 652
 FT HELIX 662 665
 FT TURN 668 669
 FT STRAND 671 678
 FT TURN 681 682
 FT STRAND 686 689
 SQ SEQUENCE 711 AA; 77101 MW; 78587D54716FC2DB CRC64;

Query Match 32.5%; Score 1155; DB 1; Length 711;
 Best Local Similarity 39.9%; Pred. No. 1e-62; 252; Idents 36; Gaps 11;
 Matches 269; Conservative 118; Mismatches 62; Polymorphisms 11;

RESULT 2
 PNP_BUCAP STANDARD; PRT; 707 AA.
 ID PNP_BUCAP
 AC Q8K9H5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) (Polynucleotide
 phosphorylase) (PNPase).
 GN PNP OR BISG361
 OS Buchnera aphidicola (subsp. *Schizaphis graminum*).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OC NCBI_TaxID:98794;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:21084349; PubMed=12089438;
 RA Tamas I.; Klasson L.; Canbaek B.; Naeelund A.K.; Eriksson A.-S.;
 RA Wernegren J.J.; Sandstrom J.P.; Moran N.A.; Andersson S.G.B.;
 RT "50 million years of genomic stasis in endosymbiotic bacteria";
 RL Science 296:2376-2379 (2002).
 CC -1- FUNCTION: Involved in RNA degradation. Hydrolyzes single-stranded
 polyribonucleotides progressively in the 3' to 5' direction.

DR SMART; SM00316; SI; 1.
 DR PROSITE; PS50084; KH_TYPE_1; 1.
 DR PROSITE; PS50126; SI; 1.
 KW Transferase; Nucleotidyltransferase; RNA-binding; Complete proteome.
 FT DOMAIN 558 592 KH.
 FT DOMAIN 623 691 SI MOTIF.
 SQ SEQUENCE 745 AA; 82339 MW; 88596545876D2F2 CRC64;
 Query Match 32.1%; Score 1140.5; DB 1; Length 745;
 Best Local Similarity 37.8%; Pred. No. 8.2e-62; Matches 236; Indels 41; Gaps 12;
 Matches 255; Conservative 143; Mismatches 236; InDel 41; Gaps 12;
 Qy 39 LEISGGKLARRPADGSAVQSGDTAVMTAVSKTKTPSPS_QMPFLVUDYKQKAAAGRPT 97
 Qy 16 LEBISTGKIAKQANAAVTVKNGNSILCTVANKYDQGFPFLPTVNYHMAKQKPG 75
 Qy 98 NYLRREVGTSRDEILTSRIDSIRPLPAGFYTQDVLQNLAVDGNEPDVLAINGAS 157
 Db 76 GPFKKGKGAKSDREILVSLRDPFLRQFMHRTVHVTCSVLSYDPAWPVILAGAS 135
 Qy 158 VALSLSDIPWNGPVGAVRIGIDGHWVNPTRKEMSSSTLNLYAGAPKSQIVMLEASAE 217
 Db 136 ALISISPAPYKELIVAKVASKVGLINGEFLVNLTELLKTSQLDLWVAGTBDS_VMVESEAH 194
 Qy 218 NILQDQFCHA1KVGVYKTOQI1QG1QQLVHETGVKTPQKFLTPSPKAVYKTHKLAMER 277
 Db 195 ILSEDKNLEAVKFGFSFOTVKLKEKAKPKPEMDQY-PS----SLIKEER 247
 Qy 278 LYAVFTDYEHDVKVSRAEVNKLRLTEBQI-----KEKPFADPYTIESRNVAKE 329
 Db 248 LFTKVEQEAEIKSKQERSTDIALYEVKUHTHFDIENKY---NNYQISEAKAISAD 304
 Qy 330 VFRSTVNLNEYKRCGDLTSLSRNVCSEVDMKFTLGSALFORQGQVQVLCVTF_DSLESG 388
 Db 305 ILRNKTELEKVRKIDGRSTTDIQRQIAEVBGVGILPSAHSALFTRGEOQSLVSTTGTSLD-- 362
 Qy 389 IKSDDVITAINGIKDKNFMAYEFPYATHEIGKOTGLANRBEIGALAKALYVPIR- 447
 Db 363 ---EQIVDSLGEYKERYKFMILNITYPPYSVNAAMPKAPSREVGHGKLMARAINPLPNK 419
 Qy 448 -DPPFTIRVTSVLENGSSMASACGSSLALMDSCVPSRSSAVAGAIGWTKTDPEKGE 506
 Db 420 WQFPVSIKVWETTESNGSSSMATVGSLLAHGVPIKAPVAGITAMGIVK-----E 472
 Qy 507 IEDYRULITDINGIEDYNGMDMPKIACTNGKGTALQADIKLPGIPIKIVMFAIQQASVAKK 566
 Db 473 SNKFAVLSDIGDDEDYFGMDPFPKAGTSSGGITALQMDKISGIDFKIQTQLEQARLGL 532
 Qy 567 EILQMKNTSKPRASKRKENGKPPVETVQVPSIKRKFVGGYNNIKLQETGVTSQD 625
 Db 533 HLEQMKVVKSPKNSLKNPSSTVVKDQKDKIKDQKIKIGPKTKEICETSNAKIDSD 592
 Qy 627 BETPSVPAFTPSVHMEARDFTBICKDQEQOLEQFGAVYTTIBRDGVNPKLYPMT 686
 Db 593 DGTVSTYIASDRDKIKIALDKIKALAEV-----EIGEIFNGTVMKVLDGAFINTVGNKD 647
 Qy 687 AVLHLNTQOLDNERLN 701
 Db 648 G-FVHISBISDARID 661

RESULT 4
 PNP_YEREN STANDARD; PRT; 706 AA.
 ID PNP_YEREN STANDARD; PRT; 706 AA.
 AC 034275; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) (Polyribonucleotide phosphorylase) (PNPase).
 DE PNP.
 OS *Versinia* enterocolitica.

OC Enterobacteriaceae; *Versinia*.
 NCBI_TAXID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98224047; PubMed=9632258;
 RA Goverde, R.J.J.; Huis in't Veld, J.H.J.; Kusters, H.G.; Mooi, F.R.;
 RT "the psychrotrophic bacterium *Versinia* enterocolitica requires expression of pnp, the gene for polyribonucleotide phosphorylase, for growth at low temperature (5 degrees C).";
 RL Microbiol. 28:555-569(1999).
 CC -1_ FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded polyribonucleotides processively in the 3' to 5' direction.
 CC Involved in the RNA degradome, a multi-enzyme complex important in RNA processing and messenger RNA degradation (BY similarity).
 CC -1_ CATALYTIC ACTIVITY: {RNA} (N+1) + phosphate = {RNA} (N) + a nucleoside diphosphate.
 CC -1_ SUBUNIT: Homodimer (BY similarity).
 CC -1_ SUBCELLULAR LOCATION: Cyttoplasmic (BY similarity).
 CC -1_ SIMILARITY: Contains 1 KH domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the SMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Y10692; CAI71697_1; ALT _INIT.
 DR HSSP; P05055; 1SD0.
 DR InterPro; IPR00147; 3_ExoNase.
 DR InterPro; IPR004187; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR003029; SI.
 DR Pfam; PF00013; KH_1.
 DR Pfam; PF03725; RNase_PH_C.
 DR Pfam; PF00575; SI; 1.
 DR SMART; SM00322; KH; 1.
 DR SMART; SM00316; KH_TYPE_1.
 DR PROSITE; PS50084; KH_TYPE_1.
 DR PROSITE; PS50126; SI; 1.
 KW Transferase; Nucleotidyltransferase; RNA-binding.
 FT DOMAIN 557 591 KH.
 FT DOMAIN 622 690 SI MOTIF.
 SQ SEQUENCE 706 AA; 76284 MW; 9CD82D348C37AD4A CRC64;
 Query Match 31.9%; Score 1136; DB 1; Length 706;
 Best Local Similarity 39.8%; Pred. No. 1.4e-61; Matches 268; Indels 32; Gaps 11; Matches 268; Conservative 115; Mismatches 259; InDel 32; Gaps 11;
 Qy 12 GQHTWTTIETGMMQATAAVWNSKDTAVFTVQGKAKPGSFFPPTVNYQERTVYAG 71
 Db 94 RIPTYLRREVGTSKELTSRIDSIRPLPAGFYTQDVLQNLAVDGNEPDVVAI 153
 Qy 72 RIPSPFPRERGRGPSEGETTSRSLRDPFLPPSFLNQVATVAVNPQINPDTIA 131
 Db 154 NGASVLSLSDIPWNGPVGAVRIGIDGHWVNPTRKEMSSSTLNLYAGAPKSQIVMLE 213
 Db 132 IGASALSLSGPFPNGPIGARVFPINDQYVNLNPDTBLKESRDLWVAGT-AGAVLME 190
 Qy 214 ASAENILQDQFCHA1KVGVYKTOQI1QG1QQLVHETGVKTPQKFLTPSPKAVYKTH-- 271
 Db 191 SEADILSEDOMLGAUVFGHQHQVVIENNALVAEG--KPKWDWHP-VNEALHAR 245
 Qy 272 --KLAMERLYAVFTDYEHDVKVSRAEVNKLRLTEBQIKEKEKPFADPYTIESRNVAKE 329
 Db 246 VAEILAAKRLGDAYRITE--KQERTVQDVKIKADTYEALIAQDDTIAQDIDGSSVKD 303
 Qy 5

CC ;-- SIMILARITY: Contains 1 KH domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL; Y18132; CMA77048.1; -.

CC HSSP; P05055; ISRO.

DR InterPro; IPR01247; 3_ExoNase.

DR InterPro; IPR04087; KH_dom.

DR InterPro; IPR004088; KH_Type_1.

DR InterPro; IPR00110; Ribosomal_S1.

DR Pfam; PF00013; KH; 1.

DR Pfam; PF03726; RNase_PH; 2.

DR Pfam; PF03725; RNase_PH_C; 2.

DR Pfam; PF00575; S1; 1.

DR PRINTS; PR00681; RIBOSOMALIS1.

DR SMART; SM00322; KH; 1.

DR SMART; SM00316; S1; 1.

DR PROSITE; PS50084; KH_TYPE_1; 1.

DR PROSITE; PSS0126; S1; 1.

KW Transferase; Nucleotidyltransferase; RNA-binding.

FT DOMAIN 558 592 KH.

FT DOMAIN 623 691 S1_MOTIF.

FT SEQUENCE 701 AA; 74984 MW; 5DBBFA446DP0F2C CRC64;

QY 35 GNRKLUSSGKLAEPAGDSAVVQ-SGDTAVMMTAVSKTKPSSQ-FNPLWYDVKQKAAA 92

QY 11 GQSTVLTETGRQATGAVLYTVNDTVLTVVQGAKQADPGKGFPLSVHQAQYTA 70

QY 93 GRIPTYLVRREGTSDKELTSRIDSIRPLEPAGYDFTQDQLNLAVDGUNEPVLA 152

QY 71 GKLPGAFPKRERGPSEGETLTSRSLIDRPIRPLPPEGFNNEVQVQUTVTSKTKDP1A 130

QY 153 INGASVALLSDP1PWNPGPVGAVRIGTIDGE-YVNNTPRKEMSSSTMILVWAGPSQIVM 211

QY 131 MIGTSALAISGIPFEGPIGAARVAFHSTGYLLNPTEQOLAASSLDNWVAGT-SDAVLM 189

QY 212 LEASAMILQDQPFCHAIIKVGWVYQTQIOTQIOLVYKETGVTQKTPQQLFTPEPEVYTH 271

QY 190 VSEADEQLETDOMLGAFLVFAHDFEBOVIAVQKELAASAG---KPTWDWKPAVANTILFN 245

QY 272 KLA MERLIVAVFTDY---EHDKVS---RDEAVNKRQLDTEBQLKERFPEADPYETES 322.

QY 246 AIRAEGRAVSGCYYTTVKADYKARGLDQAVKF---SGECQPSAS--EYKEI 297

QY 323 FNVVAKFVERFSTVLYTKEKRCBDRDLSLRNVSCEVDKPTLGSALQORGQCVLCTWTF 382

QY 298 FGRIEKRTVRENIVNGKPRIDERDNKTVRPLNTEGVYLPKTHGSALPFRGETQALVVATL 357

QY 383 DSLES5IKSDQVITAIANGIKDXNFMVHEPPYATNETGKVGTQNLNREBLGHGALAKLY 442

QY 358 GT---ARDAQDLDTLEGEKCPMPMNTNPFSVSGCGRGMGAGRBRIGHLARERSQ 413

QY 443 FVTPRO--FPTFTRTSEBVLEBENGSSMASACGSLALMDSGVPISSAVAGVAGLYVKT 500

QY 414 AMLPAADVFPYTRVVBESITESNGSSMASVCGASLALMDAGVPMKAVGAMGLVKG 473

QY 501 DPEKGEGELEYRLTDLGLIEDNGDMPFKTAGNGKITALQDILKQGIPKIKVLMVLIQ 560

QY 474 D-----KFAVLTDLGDEDHLDMDPKVAGTAGVYALQMDIKINGITEBIMELALGQ 526

DR RESULT 10

DR ECX1_SULSO STANDARD; PRT; 248 AA.

DR ID ECX1_SULSO

DR AC Q9DXC2;

DR DT 15-SEP-2003 (Rel. 42; Last sequence update)

DR DT 15-SEP-2003 (Rel. 42; Last annotation update)

DR DE Probable exosome complex exonuclease 1 (EC 3.1.13.-).

DR GN SS00735.

OS Sulfolobus solfataricus.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=2287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35092 / DSM 1617 / P2;

RC MEDLINE=20165948; Published=20101121;

RA Charlebois R.L., Singh R.K., Chan-Weiher C., Chow C.,

RA Confalonieri F., Curtiss B., Duguet M., Brausse G., Faguy D.,

RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozerka C.,

RA Kuchwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,

RA St. Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,

RA Ragan M.A., Sensen C.W.;

RT "Gene content and organization of a 281-kbp contig from the genome of the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";

RT Genome 43:116-136 (2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35092 / DSM 1617 / P2;

RC MEDLINE=21332296; Published=11427726;

RA She Q., Singh R.K., Confalonieri F., Zivarcovic Y., Allard G.,

RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtiss B.A.,

RA De Moors A., Brausse G., Fletcher C., Gordon P.M.K.,

RA Heijkamp-d' Jong I., Jeffries A.C., Kozeva C.J., Medina N., Peng X.,

RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theurault C., Tolerup N.,

RA Charlesb R.L., Doolittle R.F., Duguet M., Gaasterland T.,

RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,

RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).

CC ;-- FUNCTION: Probably involved in the 3'->5' degradation of a variety

CC of RNA species (potential).

CC ;-- SUBUNIT: Component of the archaeal exosome multienzyme

CC ribonuclease complex (potential).

CC ;-- SUBCELLULAR LOCATION: Cytoplasmic (potential).

CC ;-- SIMILARITY: BELONGS TO THE RNASE PH FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL; Y18930; CAB57560.1; -.

CC DR EMBL; AB006598; AAK4031.1; -.

CC DR PIR; H00221; H90221; -.

CC DR HAMAP; MF_00591; -; 1.

CC DR InterPro; IPR001247; 3_ExoNase.

CC Pfam; PF01138; RNase_PH; 2.

Search completed: January 8, 2004, 10:39:16
Job time : 18 secs

DR PROSITE; PS51084; KH_TYPE_1; 1.
 DR PROSITE; PS5126; S1; 1.
 KW Transferase; Nucleotidyltransferase.
 SQ SEQUENCE 783 AA; 85936 MW; 8A3629AF552F8E24 CRC64;

Query Match 98.7%; Score 312; DB 4; Length 783;
 Best Local Similarity 99.3%; Pred. No. 5.3e-219; Matches 2; Mismatches 3; Indels 0; Gaps 0;

Db 181 GEYVUNPTREMSSTNLVAGPKSQVIMLESAENILOQDCHAIKVKGKTOIQ 240
 198 GEYVUNPTREMSSTNLVAGPKSQVIMLESAENILOQDCHAIKVKGKTOIQ 257

Db 241 GIQQLVKETGVTQKLTPOKUFTPSPIVKYTHKLAMERLYAVFTDYEHKVSRAVNKR 300
 258 GIQQLVKETGVTQKLTPOKUFTPSPIVKYTHKLAMERLYAVFTDYEHKVSRAVNKR 317

Db 301 LDTEBOLKEKPEADPYTIESFNVVAKEVFRSTIVLNBYERKCDGRDLTSIRNVSCEVDMF 360
 318 LDTEBOLKEKPEADPYTIESFNVVAKEVFRSTIVLNBYERKCDGRDLTSIRNVSCEVDMF 377

Db 361 KTLGSAIQRGQVLTCTVTPDSLESGIKSQDQVITANGIKDKNFMVYEPYATNEI 420
 378 KTLGSAIQRGQVLTCTVTPDSLESGIKSQDQVITANGIKDKNFMVYEPYATNEI 437

Db 421 GRVTLNRLREBHGALAKALYVPIRDPPTIRTSEVLESNGSSMASACGSLALMD 480
 438 GKVIGLNRLREBHGALAKALYVPIRDPPTIRTSEVLESNGSSMASACGSLALMD 497

Db 481 SGVPISSAVAGVAIGLVTKDPEKEBIEYRLLTDILGIDYEDNGDMDFKLAGTNKITAL 540
 498 SGVPISSAVAGVAIGLVTKDPEKEBIEYRLLTDILGIDYEDNGDMDFKLAGTNKITAL 557

Db 541 QADIKLPGIPIKIVMEAIQASVAKKEILOMNKTKISKPRASRKEENGPPVETVQPLSR 600
 558 QADIKLPGIPIKIVMEAIQASVAKKEILOMNKTKISKPRASRKEENGPPVETVQPLSR 617

Db 601 AKFVPGGYNLKKLQAEQTVTISQVDETRSVFAPTSVHEARDPITEICKDQEQLE 660
 618 AKFVPGGYNLKKLQAEQTVTISQVDETRSVFAPTSVHEARDPITEICKDQEQLE 677

Db 661 FGAVYTATTEIRDGVWVLYKYPNTAVLHNTQLDQRKI 700
 678 FGAVYTATTEIRDGVWVLYKYPNTAVLHNTQLDQRKI 717

RESULT 2

Q8TCSS8 PRELIMINARY; PRT; 783 AA.

AC Q8TCSS8;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Polynucleotide phosphorylase-like protein (EC 2.7.7.8).
 GN PNPase.

OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]

RA SEQUENCE FROM N.A.

RA "Protein-protein interactions between human exosome components suggest
 RT the assembly of Rbse PH-type subunits into a six-membered PNPase-like
 RT ring;" Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AU458465; CAD30289; 1; -.

DR InterPro; IPR001247; 3_Exoribase.
 DR InterPro; IPR001547; Glyco_hydro_5.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_Type_1.
 DR InterPro; IPR003028; S1.

DR Pfam; PF03726; PNPase; 1.
 DR Pfam; PF01138; RNase_PH; 2.
 DR Pfam; PF03725; RNase_PH_C; 2.

DR SMART; SM0322; KH; 1.
 DR SMART; SM0316; S1; 1.

DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.

Db 601 AKFVPGGYNLKKLQAEQTVTISQVDETRSVFAPTSVHEARDPITEICKDQEQLE 660
 618 AKFVPGGYNLKKLQAEQTVTISQVDETRSVFAPTSVHEARDPITEICKDQEQLE 677

Db 661 FGAVYTATTEIRDGVWVLYKYPNTAVLHNTQLDQRKI 700
 678 FGAVYTATTEIRDGVWVLYKYPNTAVLHNTQLDQRKI 717

RESULT 3

Q8KIR3 PRELIMINARY; PRT; 783 AA.

AC Q8KIR3;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Polynucleotide phosphorylase-like protein (EC 2.7.7.8).
 GN 120003F12RIK OR PNPase.

OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.

RESULT 4						
OBR2U3	PRELIMINARY;	PRT;	589	AA.		
ID	OBR2U3					
AC	OBR2U3;					
DT	01-JUN-2002	(TREMBlre.	21	Created)		
DT	01-JUN-2002	(TREMBlre.	21	Last sequence update)		
DT	01-MAR-2003	(TREMBlre.	23	Last annotation update)		
DE	Similar to putative.					
GN	1200003P12RICKN					
OS	Mus musculus (Mouse)					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus					
OX	NCBI_TaxID:10090;					
RN	[1] SEQUNENCE FROM N.A.					
RA	Strauberg R.;					
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; BC027228; AAH27228.1; -;					
MGI	MGI:1918951; 1200003P12RICK.					
DR	InterPro; IPR01247; 3_EORNASE.					
DR	InterPro; IPR04087; KH_dom.					
DR	InterPro; IPR00488; KH_type_1.					
DR	InterPro; IPR01029; SL_.					
DR	PFam; PF03726; RNase; PH; 1.					
DR	PFam; PF01138; RNase; PH; 1.					
DR	SMART; SM00322; KH; 1.					
DR	SMART; SM00316; SL; 1.					
DR	PROSITE; PS50084; KH_TYPE_1; 1.					
DR	PROSITE; PS50126; KH; 1.					
DR	SEQUNCE; 589 AA; 6496 MW; 808162BB46F6428F CRC64;					
Query	Match	69.8%	Score	2484	DB	11; Length 589;
Best	Local Similarity	92.2%	Pred.	No.	1	4e-152;
Matches	482;	Conservative	26;	Mismatches	15;	Indels 0;
Matches						Gaps 0;
QY	178 IIDGEVYVNPTRKEMMSSTLNVWAGARKPSQINMLASAENTIQQDCHAIVKGVVKYQQ	237				
QY	1 MIDGECTVNPTREMSSTLNVWAGARKPSQINMLASAENTIQQDCHAIVKGVVKYQQ	60				
Db	238 IIGQIQLVKGKGVTRTPQKFTPSBPIVKYTHKLMERLYAVFTYEHDKVSRDEAVN	297				
Db	61 IIGQIQLVKGKGVTRTPQKFTPSBPIVKYTKIANKELKAVFTYEHDKVSRDEAVN	120				
QY	298 KIRLDTEBOLKEKPPADPYETIESAVVAKENFRSIVINVEYKRCGDRDTSLRNSCEV	357				
Db	121 KIRLDTEBOLKEKPPADPYETIESAVVAKENFRSIVINVEYKRCGDRDTSLRNSCEV	180				
QY	358 DMFTKLIGSALFORGTVOTVLTFTFDSLESGTSKSDQVITAIINGIKDNFMILYEPPTAT	417				
Db	181 DMFTKLIGSALFORGTVOTVLTFTFDSLESGTSKSDQVITAIINGIKDNFMILYEPPTAT	240				
QY	418 NEIGKVGVLNRELHGIGALAKALYPTIPRDPFTTRTVTSELESGSSSSMASACGSSLA	477				
Db	241 NEIGKVGVLNRELHGIGALAKALYPTIPRDPFTTRTVTSELESGSSSSMASACGSSLA	300				
QY	478 LMDSGWVIISSAVAGVAGVAGLVIKTDPEKGIEDYRLLTDIGBEDNGMDFKTAGTNKG	537				
Db	301 LMDAGVPISSAVAGVAGLVIKTNPENGEGRIEDYRLLTDIGBEDNGMDFKAGTNKG	360				
QY	538 TALQADIKLPGPIKTYMEAOQASVARKETIQLIMKNTISKPRASKRKGKENGPPVETVQPL	597				
Db	361 TALQADIKLPGPIKTYMEAOQASVARKETIQLIMKNTISKPRASKRKGKENGPPVETVQPL	420				
QY	598 SRAKAKFVGPGCGNKLQKQAOEGFTISVQDDEPSVAFPTPSWHRADFTEICKDQEQ	657				
Db	421 SRAKAKFVGPGCGNKLQKQAOEGFTISVQDDEPSVAFPTPSWHRADFTEICKDQEQ	480				

QY	658	QLEFGAVVTTITERDTGIVNKLYPMNTAVLANTQDNERL	700	Db	259	IQQLVKEISVAKRTPQK1KFTPSABIVKTKLAVFTDYEHDKVSRDAVKTRL	318	
Db	481	QLEFGAVVTTITERDTGIVNKLYPMNTAVLANTQDNERL	523					
RESULT 5								
Q9DC52		PRELIMINARY;	PRT;	540 AA.				
ID	Q9DC52							
AC	Q9DC52;							
DT	01-JUN-2001	(TREMBrel. 17, Created)						
DT	01-JUN-2001	(TREMBrel. 17, Last sequence update)						
DT	01-MAR-2003	(TREMBrel. 23, Last annotation update)						
DB	120003F12R1K.							
GN	Mus musculus (Mouse)							
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; OC							
OC	NCBI_TaxID=10090;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC								
RX	MEDLINE=>2108560; PubMed=11217851;							
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Furukoshi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batyalov S., Casavant T., Fleischmann W., Gasterstland T., Glass C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schirali L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gusarincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Maszarek J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Suzuki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-P., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynnshaw-Boriss A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;							
RT	"Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).							
DR	EMBL; AK004563; BAB23374_1; -;							
DR	MGD; MGI:191851; 1200331215k.							
DR	InterPro; IPR001247; 3_ExonRNAse.							
DR	InterPro; IPR001547; Glyco_hydro_5.							
DR	PFAM; PF03726; PNPase; 1.							
DR	PFAM; PF01138; RNase_PH; 2.							
DR	PFAM; PR03725; RNase_PH; 1.							
DR	PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.							
DR	SEQUENCE; PS00659; GLYCOSYL_HYDROL_F5; 1.							
DR	540 AA; 58938 MW; 457BFA3B3579A072 CRC64;							
SQ								
Query Match	67.2%; Score 2389; DB 11; Length 540; Best Local Similarity 90.1%; Pred. No. 1.8e-146; Matches 465; Conservative 29; Mismatches 22; Indels 0; Gaps 0;							
Query	2 GPFLLRRDRALTQVQRALWSSAGSRAVADGKRLTSKGKLFADGGAVVQSGT	61	Query Match	64.2%; Score 2285; DB 4; Length 504; Best Local Similarity 100.0%; Pred. No. 8.9e-10; Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Db	19 GPLGRGRGRNRLAISYLOMRAJISSTSSRAVTDGRKLEISSKGKLFADGGAVQSGT	78	Query	1 DGPFLPDRDRALTQVQRALWSSAGSRAVADGKRLTSKGKLFADGGAVVQSGD	60	Query	61 TAVMVTAVSKTKPSQSPQMLPVYDQKAAAGRIPNYLRLREVTSKEITSRISI	120
QY	62 AVMVTAVSKTKPSQSPQMLPVYDQKAAAGRIPNYLRLREVTSKEITSRISI	121	Db	18 DGPFLPDRDRALTQVQRALWSSAGSRAVADGKRLTSKGKLFADGGAVVQSGD	77	QY	61 TAVMVTAVSKTKPSQSPQMLPVYDQKAAAGRIPNYLRLREVTSKEITSRISI	130
Db	79 AVMVTAVSKTKPSQSPQMLPVYDQKAAAGRIPNYLRLREVTSKEITSRISI	138	Db	78 TAVMVTAVSKTKPSQSPQMLPVYDQKAAAGRIPNYLRLREVTSKEITSRISI	137	QY	121 IRPLPAGYFVDTQVLCNLALVDGNEPDVLAINGASVIALSLSDIPWNPVGA	180
QY	122 RPLPAGYFVDTQVLCNLALVDGNEPDVLAINGASVIALSLSDIPWNPVGA	181	QY	121 IRPLPAGYFVDTQVLCNLALVDGNEPDVLAINGASVIALSLSDIPWNPVGA	198	QY	138 IRPLPAGYFVDTQVLCNLALVDGNEPDVLAINGASVIALSLSDIPWNPVGA	197
Db	139 RPLPAGYFVDTQVLCNLALVDGNEPDVLAINGASVIALSLSDIPWNPVGA	198	Db	139 RPLPAGYFVDTQVLCNLALVDGNEPDVLAINGASVIALSLSDIPWNPVGA	258	QY	181 GEYVNPNTRKEMSSSTLNVWAGPKSQIVMELASAEQNLQDFCHAIVKGVKTQ	240
QY	182 GEYVNPNTRKEMSSSTLNVWAGPKSQIVMELASAEQNLQDFCHAIVKGVKTQ	241	Db	198 GEYVNPNTRKEMSSSTLNVWAGPKSQIVMELASAEQNLQDFCHAIVKGVKTQ	257	QY	198 GEYVNPNTRKEMSSSTLNVWAGPKSQIVMELASAEQNLQDFCHAIVKGVKTQ	301
QY	242 IQLQVKEISVAKRTPQK1KFTPSABIVKTKLAVFTDYEHDKVSRDAVKTRL	301						

RESULT 11
 Q965N3 PRELIMINARY; PRT; 745 AA.
 ID Q965N3; AC DT 01-DEC-2001 (TREMBrel. 19, Created)
 AC DT 01-MAR-2002 (TREMBrel. 20, Last sequence update)
 AC DT 01-MAR-2003 (TREMBrel. 23, Last annotation update)
 DE Hypothetical 82.9 kDa protein.
 GN BE0003N10.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderrinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1] SEQUENCE FROM N.A.
 RP RX STRAIN=Bristol N2;
 RP RC MEDLINE=99069613; PubMed=9851916;
 RA RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2] SEQUENCE FROM N.A.
 RP RC STRAIN=Bristol N2;
 RA RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC092690; AAKEY3855.2; -.
 DR InterPro; IPR001247; 3_ExonNase.
 DR Pfam; PF03726; PRNase; 1.
 DR Pfam; PF01138; RNase_PH; 2.
 DR Pfam; PF03725; RNase_PH_C; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 745 AA; 82881 MW; 45D03070C72ACB16 CRC64;
 Query Match 35.9%; Score 1276; DB 5; Length 745;
 Best Local Similarity 42.2%; Pred. No. 4.1e-74;
 Matches 278; Conservative 109; Mismatches 237; Indels 34; Gaps 10;
 DR 28 RAVAVDL-GKVKLRSISGKLAKEADGSAVQSGDTAVMTAVS-KTKRSQPSQFVPLVYD 85
 Db 26 QSVAVDFDGKSGKIELKAGHLRFASTVVASFGDNAMAVTVORKSKQGDGSVPLQVEY 85
 QY 86 RQKAAAGRPTVNLREVTSKELTSRIDIIDSIRPLPPAGFVYDQVLCNLAVDGV 145
 DR 146 NEPDVLAINGASVAVSLSDIFWNGVGAWRIG-IIDGEYVNVPRKEMSSSTNLVAGA 204
 QY 145 ADQIMGGINAATQIQQSSAVYNGLALAVRVARPARGDFHVNPIQEBELREASNLIVAR 204
 Db 205 PKSQVUMLESAENILOQDFCHAIKVGKTTQIIGQIQLVKEVGTGVTKRPQLFTSP 264
 DR 205 KHEKTVIEMLDGREGSAEHLHALDVAFRVAKLHEAMQLTAR-----PKD 251
 QY 265 EIV-----KTHIKLAMERRLYAVFTDYEHDKYSRDEAVNKIRLDETEQLEKEK-PPEAD 315
 DR 252 ELASEDPGSLERLJBETARERIYVYITDAGHKLISRDMEIKAI---FEBICAQFQCE 308
 QY 316 PYEIESFNVVAKEFVRFSTVUNVEYKRGCDGRDLTSIRANVSCBVDMFKTLHGSALFORQTO 375
 Db 309 KDAIYRTSYSTVKKLDRDTLRTGCRPDRPFRPITHDVKKUHGCSIFQRQGQ 368
 QY 376 VLCTVYFDSIIESGKSDQVTAINGIKDKNFMAYEFPYPAATNEIGKVGVLNRLRBLGIGA 435
 Db 369 WSTMVTFDSPAFAFPHDPSVQQLGSQRKSFMLHVEPGPFAINEFGTTSRSLNRBIGHCA 428
 QY 436 LAKKALYVPIRDPFPTITVSEVLESNSSSMSMASACGSLSALMDSCVPIISAVAGVATG 435
 Db 429 LAEKSILKNLFPADFPYATRLACQVLESNGSSMSMASVCGSSLALFDAGVPMKAAGVAIG 488

RESULT 12
 Q9RA43 PRELIMINARY; PRT; 707 AA.
 ID Q9RA43; AC DT 01-JUN-2002 (TREMBrel. 21, Created)
 AC DT 01-JUN-2002 (TREMBrel. 21, Last sequence update)
 AC DT 01-MAR-2003 (TREMBrel. 23, Last annotation update)
 DE Polyribonucleotide nucleotidyltransferase (Polynucleotide DE phosphorylase).
 GN PNP OR TTE187.
 OS Thermomaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermomaerobacterales;
 OC Thermomaerobacteriaceae; Thermomaerobacter.
 OC NCBI_TaxID=119072;
 RN [1] SEQUENCE FROM N.A.
 RP RC STRAIN=MB4 / JCM 11007;
 RX MEDLINE=2192816; PubMed=1199736;
 RA Bao Q., Tian Y., Li W., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yang J., Yang H.,
 RT "A complete sequence of *T. tengcongensis* genome.";
 RL Genome Res. 12:659-700(2002).
 DR EMBL; AE013098; NM24609.1; -.
 DR InterPro; IPR001247; 3_ExonNase.
 DR InterPro; IPR00487; KH dom.
 DR InterPro; IPR00488; KH_type_1.
 DR InterPro; IPR00329; SI.
 DR Pfam; PF00013; KH; 1.
 DR Pfam; PF03726; PRNase; 1.
 DR Pfam; PF01138; RNase_PH; 2.
 DR Pfam; PF03725; RNase_PH_C; 2.
 DR Pfam; PF00575; SI; 1.
 DR SMART; SM00322; KH; 1.
 DR SMART; SM00316; SI; 1.
 DR PROSITE; PS00084; KH_TYPE_1; 1.
 DR PROSITE; PS50126; SI; 1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 707 AA; 78519 MW; 779E64C04BBA3331 CRC64;
 Query Match 35.2%; Score 1253.5; DB 16; Length 707;
 Best Local Similarity 40.0%; Pred. No. 1.1e-72;
 Matches 275; Conservative 135; Mismatches 230; Indels 47; Gaps 13;
 DR 28 RAVAVDL-GKVKLRSISGKLAKEADGSAVQSGDTAVMTAVSKTKPSBS-QFVPLVYD 86
 Db 4 RTFEMELAGRKLVVOTGKVAQGAAWVKGVTQIQLVKTGVTQKPLFTSP 63
 QY 87 QKAAAGRPTVNLREVTSKELTSRIDIIDSIRPLPPAGFVYDQVLCNLAVDGV 146
 Db 64 ERULYSIGKPGFTRKSKPSEKALSLRRLPRLPFCYRNDVQVIAWLVSPDDA 123
 QY 147 EVDVLTAVGASVAVSLSDIFWNGVGAWRIG-IIDGEYVNVPRKEMSSSTNLVAGPK 206
 Db 124 QPEIYAMIGGSSVAVLISIDIPNPGPTGSVAVGLVDKFVNPYBREKSLMLHVSGT-K 182
 QY 207 SQVUMLESAENILOQDFCHAIKVGKTTQIIGQIQLVKEVGTGVTQKPLFTSP 266
 Db 183 DATVNEAGAKEVPEETTIDAIYVHAQYIKQIVEFIRGIVKEVGV-----PKREV 232

SO	SEQUENCE	703 AA;	77989 MW;	5EBC83EDEDCA2B CRC64;
QY	267 VCKTHKLMF---RUYAVFDYE---HOKVSRDAVKIRLDRTEBOLKEKFPF--AD	315		
Db	233 V---LHRIDKELERKVKRAYATEKLYNALRTPPEKKERNDLKVYEQVLEHPEPONLAD	290		
QY	316 PWEISSFNVVAKERFERSIVLNBVKRCDCGRDILTSRANVSCEDMFKTLHGSALFORQOTO	375		
Db	291 IDEVLUK---IMKEQMRKMKMKKIRVGDGRDIFRIPWCEVGVLPRTHGSAIFTRGQTO	347		
QY	376 VICTVTFDSLSESGIKSKPDVITANGIDKDNFLHYEPPYANNEIGVTKGTRRELGHA	435		
Db	348 VITVATL---GAIGDQIILEGIGDEERKRYMHNPPYSGEVRLPGRGGRRELGHA	403		
QY	436 LAKKALYVPIP--RDPFTIRTSEVLESLNGSSMASACGSSLALMDSGVVISAVAGVA	493		
Db	404 LAERALEPVPISBEEFPPTIRSEVLESLNGSSTSQASCSTALADAGVPIKAPYAGIA	463		
QY	494 IGLVTKTDPEKOBIEDYRLLTDIGTIDYNGMDKFTAGTNGKITALQADIKLPGIPIK	553		
Db	464 MGLIKED-----EVILITDQIGBEDFLGMDMDFKVAGTKEGVTLQMDIKIPGIDRDI	516		
QY	554 VNEA1QOASVAKKELQIMKTSKPKRSRCKENGPPVETVQPLSRAKFKVPGGYNLKK	613		
Db	517 LOMALEKARKARKYLVQKMLEVKEPKRKLSEKPKYAPTRVWNPBKIRDIGPAGKMTK	576		
QY	614 LOAETGTTISQDBETFSVPAFTPSVMEHARDPITEICKDQEQQLEPFGAVTATTEIR	673		
Db	577 IISETYKIDIDEDGRYVITATNLEAGERAKQMIATAKTD---IEVGGTYLGKVLIA	631		
QY	674 DTGVMVTKLYPNTMATTAVLNTQDNERL 700			
Db	632 PFGAFVBEIAPKCKEG-LVHISNLSSKKRV 657			
RESULT 13				
Q97145	PRELIMINARY; PRT; 703 AA.			
AC	097145;			
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)			
DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)			
DB	Polyribonucleotide nucleotidyltransferase.			
GN	CAC108.			
OS	Clostridium acetobutylicum.			
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OX	Clostridium.			
RN	[1] NCBI_TaxID:1488;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 824 / DSM 792 / VKM B-1787;			
RX	MEDLINE=21359325; PubMed=1146286;			
RA	Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J.J., Qiu D., Hitti J., Wolf Y.I., Daly M.J., Tatusov R.I., Sabat P., Doucette-Stamm L., Soucaille P.,			
RA	Bennett G.N., Koonin E.V., Smith D.R.;			
RT	"Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum.",			
RL	J. Bacteriol. 183:4323-4338 (2001).			
DR	EMBL; AE007689; ARK9773.1; -;			
DR	InterPro; IPR001247; 3_ExorNase.			
DR	InterPro; IPR004087; KH_dom.			
DR	InterPro; IPR00408; KH_type_1.			
DR	InterPro; IPR003029; S1.			
DR	Pfam; PF00013; KH_1.			
DR	PF03726; RNase; 1.			
DR	Pfam; PF01138; RNase; PH_2.			
DR	Pfam; PF03725; RNase_PH_C; 2.			
DR	Pfam; PF00575; S1; 1.			
DR	SMART; SM00322; KH; 1.			
DR	PROSITE; PS50084; KH_TYPE_1; 1.			
DR	PROSITE; PS50126; S1; 1.			
KW	Transferase; Complete proteome.			
RESULT 14				
Q8KBV3	PRELIMINARY; PRT; 733 AA.			
ID	Q8KBV3;			
AC	Q8KBV3;			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DB	Polyribonucleotide nucleotidyltransferase.			
GN	CNP OR Ctn649.			
OS	Chlorobium tepidum.			
OC	Bacteria; Chlorobi; Chlorobiales; Chlorobiaceae;			
OX	Chlorobium.			
RN	NCBI_TaxID:1097;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TLS / ATCC 49652 / DSM 12025;			
RX	MEDLINE=2110385; PubMed=1203901;			
RA	Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,			

QY 257 QKLFPTPS-----PEVKYTHKLAMERLYAVFTDYEDKVERDEANKIRDTEQIKE 309
 Db 233 ---FIPAEERDEALVERVKSLLTSEKGKETVLT-----DKOORDENIDNLK---BEI VNE 281
 QY 310 KFPEADP-----YELESFNVVAKEFVRSIVLANEYKRCGDLTSLANSCEVDMKTL 363
 Db 282 FIDEEDPENELIJKEVTAILENLVEKETRLLADRKIRPDPGRKPDETRPLOSEVGLPRT 341
 QY 364 HESALFORGQOTQVLCVTFDSIERSGIKSQDVTAINGKDKNFMALYTFPPVATNEIGV 423
 Db 342 HSSGLFTRGQTOQALSVLTIGAL---GDIQLINGLGPEEKRFMHHTNFPNPSVGEGP 397
 QY 424 TGINRRELGHGALAEKALYVIP---RDPPTTRVTSEVLESUNGSSMASACCGSLALMS 481
 Db 398 RAPGRREBIGHGALGERALKYTFDTADPPTTRIVSEYLESNGSSOASICGSTLALMDA 457
 QY 482 GPISSAVAGVAIGLVTKDPEKEIEDYRLLTDLGEDBDMDFKIAGTNKGITALQ 541
 Db 458 GYPIKAPVAGIANGLVTRD-----SYTILTDIQMEDALGMDPKVASJKEGITAQ 510
 QY 542 ADTKLPGJPIKTMETAOQASVAKBETIQMINKTIISKPRASRKENGPVWETVOPLSKA 601
 Db 511 MDIKIDGLTRE-LBEALEQARRGRLEINNMHLQTDQPTBLTAYPKVUTWTKPDKIR 570
 QY 602 KERVPGGYNKLQOAEGTUTISQVDEEFSPVAPTMHEARDPITRICKDQEQOLEP 661
 Db 571 DVIGPGKXKINELIDETGKVLDEQDGTTIFGAVDQMINRABEELBEITR----BAEV 625
 QY 662 GAVTATITEIRDITGVWKLYPNTAVLHNTOLDNRL 700
 Db 626 GQRYQATVKRIEKYGAFFVGLFFCKDA-LJHISQISKUR 663

Search completed: January 8, 2004, 10:40:18
 Job time : 50 secs

